Evaluating the Joint Implementation of Congestion Pricing and Driver Information Systems

by

Ibrahim Mahmoud El Sanhouri

B.S., University of Khartoum (1986)
S.M., Massachusetts Institute of Technology (1990)

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1994

© Massachusetts Institute of Technology 1994. All rights reserved.

Authored by

Department of Civil and Environmental Engineering
February 8, 1994

Certified by

Professor David H. Bernstein
Thesis Supervisor

Accepted by

Professor Joseph Sussman
Chairman, Departmental Committee on Graduate Students

MAR 21 1994
ARCHIVES
Evaluating the Joint Implementation of Congestion Pricing and Driver Information Systems

by

Ibrahim Mahmoud El Sanhouri

Submitted to the Department of Civil and Environmental Engineering on February 8, 1994, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

In this thesis, the joint implementation of congestion pricing and driver information systems is analyzed. After selecting specific pricing and information systems from amongst the many possible choices based on representativeness and implementability, a modeling framework is developed that analyzes the interaction of the two systems in the presence of incidents, or short-term reductions in network capacity. First the impact of toll station is examined by looking at the benefits and costs when stations are either upstream or downstream of the points of congestion. It is found that downstream toll stations distort the arrival times at the toll stations, resulting in toll costs that may be hard to predict. In the upstream case, negative net benefits in costs including tolls were occasionally observed. This occurred when the information system increased the total number of toll paying travelers. Modifying the information systems through reducing guidance rates eliminated or reduced this problem. Two modified pricing systems were also examined. The first consisted of tolls that were identical across routes for better implementability. A side effect was that revenue-increasing switching was eliminated and negative benefits no longer observed. In the second, only one route was tolled. In that case the optimal route split may not be attainable because ideally we would like more travelers on the tolled route than is possible with the restriction on the untolled route. Finally, superadditivity was examined and it is found that for our model, only simple additivity holds.

Thesis Supervisor: Professor David H. Bernstein
Title: Assistant Professor of Civil and Environmental Engineering
To Amani
In the Name of Allah, The Most Merciful, the Most Compassionate

Thanks To:

Professor David Bernstein, the supervisor of this dissertation, who provided ideas, insight, guidance, and encouragement in abundance. I am especially grateful for his willingness throughout these years to spend literally hours to discuss all aspects of this research (even if it meant he would get home late!).

Professors Nigel Wilson and Moshe Ben-Akiva, the members of my thesis committee, who have helped to improve the quality of the thesis with numerous suggestions, and who have provided encouragement during my entire graduate stay at MIT.

Professor Yossi Sheffi, for his support during the final phase.

Anil Mukundan, designer of the “MIT simulator” used in the research, for many hours of help regarding its use.

The many friends I have made here, both in the department and at the MIT masjid. Too numerous to list all, I must mention two for their long “service”: El Fatih El Tahir, a most reliable friend, and Arif Khalid, who is truly like a brother to me. And office-mate Dale Lewis, for his time as the clock was winding down.

My parents, Helga and Mahmoud, my sister Mona, and my brother Fawzi, for moral and material support.

And of course, my wife Amani, who was with me all the way, for her love and patience during this long stretch, and my son Ahmed, for making us happy.

All Praise be to Allah, the Lord of the Worlds.
Contents

1 Introduction
   1.1 The Congestion Problem .............................................. 20
   1.2 Traditional Approaches of Dealing with the Problem ............. 23
       1.2.1 Supply Augmentation ........................................ 23
       1.2.2 Demand Management .......................................... 24
   1.3 Innovative Approaches to the Congestion Problem .............. 25
   1.4 Research Objectives ................................................ 28
   1.5 Thesis Outline ....................................................... 29

2 Previous Research
   2.1 Congestion Pricing ................................................... 31
       2.1.1 Theoretical Literature ...................................... 31
       2.1.2 Practical Applications of Congestion Pricing ............. 38
   2.2 Advanced Traveler Information Systems ............................ 42
       2.2.1 Practical Applications ...................................... 43
       2.2.2 Theoretical Work ............................................. 44
   2.3 Combined Congestion Pricing / Traveler Information Systems .... 46

3 Identifying Candidate Combined Systems ............................... 49
   3.1 Characterization of Congestion Pricing and Traveler Information Systems ........................................ 50
       3.1.1 Characterization of Congestion Pricing Schemes ............ 50
       3.1.2 Characterization of Traveler Information Systems ......... 56
   3.2 Desirable Properties of Pricing and Information Schemes ...... 58
       3.2.1 Congestion Pricing ............................................. 58
       3.2.2 Information Systems ........................................... 61
   3.3 Selecting A Congestion Pricing Scheme ...................... 62
       3.3.1 Temporal Structure ........................................... 63
       3.3.2 Spatial Structure: Zone vs Facility Pricing .............. 66
       3.3.3 Selecting a Scheme .......................................... 67
List of Figures

1-3 Total Highway Fuel Consumption Source: US DOT (1990) ...................... 23

2-1 Static Models (Source: Small [1992c]) .................................................. 33
2-2 A two-route network with inelastic total demand ..................................... 34
2-3 The theory: Charging the difference between social and private cost results in system optimum ......................................................... 35

3-1 Congestion Pricing Design Factors ......................................................... 51
3-2 Interval Based Tolls: A Single-Step Toll ................................................. 53
3-3 Examples of Zone-Based Pricing .......................................................... 55
3-4 Information System Design Factors ...................................................... 57

4-1 The Network ......................................................................................... 78

5-1 Equilibrium costs for untolled regime ...................................................... 91
5-2 Equilibrium costs for toll regime (excl. tolls) ........................................ 92
5-3 Equilibrium costs for toll regime (incl. tolls) .......................................... 92
5-4 Costs with incident on route 1 - Untolled regime, No Information .......... 94
5-5 Costs with incident on route 1 - Untolled Regime, With Information ...... 95
5-6 Costs with incident on route 2 - Untolled Regime, No Information ......... 95
5-7 Costs with incident on route 2 - Untolled Regime, With Information ...... 96
5-8 Benefitogram - Incident on route 1, Untolled Regime ........................... 98
5-9 Benefitogram - Incident on route 2, Untolled Regime ........................... 98
5-10 Average Improvements for various guidance rates ............................... 99
5-11 Benefitogram - Incident on route 1, 25% Info, Untolled Regime .......... 99
5-12 Histogram of benefits - Untolled regime .............................................. 101
5-13 Costs (excl. tolls) - Incident on Route 1, No Info, Upstream Toll Station .. 104
LIST OF FIGURES

5-14 Costs (excl. tolls) - Incident on Route 1, With Info, Upstream Toll Station . 105
5-15 Benefitogram (excl. tolls) - Incident on Route 1, Upstream Toll Station . 105
5-16 Costs (incl. tolls) - Incident on Route 1, No Info, Upstream Toll Station . 106
5-17 Costs (incl. tolls) - Incident on Route 1, With Info, Upstream Toll Station . 106
5-18 Benefitogram (incl. tolls) - Incident on Route 1, Upstream Toll Station . 107
5-19 Benefits for Upstream Locations: Incident on Route 1 . 107
5-20 Benefits for Upstream Locations: Incident on Route 2 . 108
5-21 Number of Toll Paying Commuters with Information: Incident on 1 . 108
5-22 Number of Toll Paying Commuters with Information: Incident on 2 . 109
5-23 Costs (excl. tolls) - Incident on Route 1, No Info, Downstream Toll Station . 111
5-24 Cost (excl. tolls) - Incident on Route 1, With Info, Downstream Toll Station . 112
5-25 Benefitogram (excl. tolls) - Incident on Route 1, Downstream Toll Station . 112
5-26 Costs (incl. tolls) - Incident on Route 1, No Info, Downstream Toll Station . 113
5-27 Costs (incl. tolls) - Incident on Route 1, With Info, Downstream Toll Station . 113
5-28 Benefitogram (incl. tolls) - Incident on Route 1, Downstream Toll Station . 113
5-29 Toll Station Location and Cost incl. tolls: No Info . 114
5-30 Toll Station Location and Cost incl. tolls: Info . 115
5-31 Benefitogram (incl. tolls) - Upstream Toll Station, 25 % Information . 116
5-32 Benefitogram (incl. tolls) - Upstream Toll Station, No Info for Bulks . 117
5-33 Bulk-off and 25% guidance compared to 100% guidance: Incl. Tolls, Upstream Toll Station . 119
5-34 Bulk-off and 25% guidance compared to 100% guidance: Excl. Tolls, Upstream Toll Station . 120
5-35 Bulk-off and 25% guidance compared to 100% guidance: Incl. Tolls, Downstream Toll Station . 123
5-36 Bulk-off and 25% guidance compared to 100% guidance: Excl. Tolls, Downstream Toll Station . 124
5-37 Costs (incl. tolls) - Info, Upstream IR Tolls . 128
5-38 Benefitogram (incl. tolls) - Upstream IR Tolls . 128
5-39 Improvement over No-Info (incl.tolls) for IR Tolls, Upstream . 129
5-40 Benefitogram (incl. tolls) - Downstream IR Tolls . 129
5-41 Improvement over No-Info (incl.tolls) for IR Tolls, Downstream . 130
5-42 Total Costs as a function of flow on tolled route 1 . 133
5-43 Costs (incl. tolls) - No Info, Toll on Route 1, Incident on 1 . 135
5-44 Costs (incl. tolls) - Info, Toll on Route 1, Incident on 1 . 135
5-45 Benefitogram (incl. tolls) - Toll on Route 1, Incident on 1 . 136
5-46 Benefitogram (incl. tolls) - Upstream Toll on Route 1 only, Incident on 2 . 136
5-47 Improvements in Costs (incl. tolls) Incident on 1 . 137
LIST OF FIGURES

5-48 Improvements in Costs (incl. tolls) Incident on 2 .................................. 137
5-49 Evaluating Superadditivity: Incident on route 1 .................................... 140
5-50 Evaluating Superadditivity: Incident on route 1, Component Curves ........ 140
5-51 Evaluating Superadditivity: Incident on route 2 .................................... 141
5-52 Evaluating Superadditivity: Incident on route 2, Component Curves ........ 141
5-53 Evaluating Superadditivity: Incident on route 1, No-Toll Route Splits ...... 142
5-54 Evaluating Superadditivity: Incident on route 2, No-Toll Route Splits ...... 142
List of Tables

4.1 Specific Integrated Design Issues ........................................ 74
5.1 No-Toll equilibrium and equilibrium with optimal step-toll .......... 93
5.2 Means for Various Combinations of Toll Station Location and Info Systems for a Particular Incident ................................. 117
5.3 Summary Statistics for Various Systems: Upstream .................. 118
5.4 Summary Statistics for Various Systems: Downstream ............... 122
5.5 Optimal tolls for identical and reasonable toll structures ........... 126
5.6 Identical and Reasonable Toll Structures: Statistics for an Example 126
5.7 Optimal route splits ignoring uniform toll constraint ............... 131
5.8 Equilibrium Route Splits with zero uniform toll on tolled route .... 132
Chapter 1

Introduction

Many industries charge a price for their products that depends on when they are purchased. Placing a phone call is cheaper in the evening than during busy work hours. Electricity rates are lowest overnight. During the off-season, hotel fares are often dramatically lower than during the time when demand is highest. The purpose of such pricing policies is to better utilize facilities and/or mitigate the ill-effects of intense loads.

Excessive traffic congestion, seen in cities around the world, attests to the fact that there is a similar “problem” with the demand for highway transportation. Using pricing to improve matters by affecting traveler behavior is called congestion pricing, and forms the main topic of the research presented in this thesis.

Congestion pricing is a powerful tool for modifying travel demand patterns and reducing congestion. Pricing can obviously be used in a variety of ways to affect travel demand. For example, we can price cars by adding a surplus charge to the purchase price or through registration fees\(^1\), or through insurance premia. Pricing methods that more directly impact travel demand include gasoline taxes and parking charges. Congestion pricing, however, directly seeks to modify, through pricing, certain patterns of car use that are particularly inefficient, such as travel under highly congested conditions.

Although the theory behind congestion pricing has received a great deal of attention, it is only now that increasing effort is being spent on its practical implementation. At the same time, traveler information systems are being tested extensively in several countries. Their objective is to help commuters make better travel decisions. The issues arising in the joint implementation of pricing and information systems are investigated in this thesis.

This chapter motivates the research through a discussion of the congestion problem and the tools used in alleviating it. The main research objectives are then outlined, and the

---

\(^1\)These techniques are more directly aimed at car ownership. Of course, the demand for car ownership may exhibit some elasticity with regard to the price of car use at congested conditions, and this may in turn affect car use even under non-congested conditions. See Altshuler et al. (1979).
chapter concludes by laying out the structure of the remainder of the thesis.

1.1 The Congestion Problem

There is a widespread belief that “something must be done” about highway congestion. The problem has been receiving much publicity, such as cover page stories in news magazines (e.g. Time [1990], Economist [1991]). There are several reasons for this. The most important is that congestion, as measured in vehicle-hours “wasted”, has increased. In the U.S., one reason for this increase is that, as two-income families have become more prevalent and suburban rings have grown in both absolute population and area, the number of commuters using private automobiles has increased dramatically. Thus from 1960 to 1980 the number of people driving to work nearly doubled, whereas the number of people in the workforce increased by only 50% (Meyer et al. [1989]). Figure 1-1 shows that the demand for automobiles continues to grow at a faster rate than the population. Moreover, Figures 1-2 and 1-3 show that total vehicle miles traveled, and highway use of gasoline, have also grown steadily, despite some brief periods affected by the petroleum shortfalls in 1973 and 1978 (US DOT [1990]).

The economic costs of this congestion are believed to be huge, and projections for future costs suggest that matters will only get worse2 (Meyer et al. [1989]). Thus, excessive energy consumption due to congestion by automobiles is a concern in its own right due to its economic and political implications.

Unfortunately, there are other facets of this problem. Some of these are not just a function of congestion because they are expected to decrease but not to vanish entirely as congestion is reduced. An important social problem which is only indirectly caused by congestion is poor highway safety. Annual highway accident fatalities remain above the 40,000 mark in the U.S. Several studies have also analyzed the effect of driving under highly congested conditions on health, such as the rise in blood pressure (Stokols and Novaco [1981]). The most visible of the indirect aspects of congestion today is environmental damage, which is discussed below.

Environmental concerns form one of the major forces behind calls for measures to inhibit congestion. In that regard it must be observed that the primary legislative document, the Clean Air Act of 1970, underwent comprehensive revision in 1990. A great deal of the specific regulations are still in the process of negotiation and finalization, but a substantial part deals with transportation in urban areas. In particular, stringent requirements are specified on various aspects of travel, such as ridesharing levels. Some of the reasons for

---

2 Estimates of the economic costs of congestion are not very precise, though. Most studies assume that the value of travel time of an individual is constant even for very small time periods, an assumption that tends to overestimate those costs.
the increased emphasis on transportation are (Salvucci [1993]):

- The evolving body of knowledge on the environmental impacts of transportation related pollutants has confirmed the worst expectations of only a short time ago. An example in point is the case of the combustion by-product carbon dioxide, which has recently been linked to the erosion of the protective ozone layer in the atmosphere, but was not previously considered a major pollutant.

- The transportation sector has lagged in “cleaning up its act” compared with other sectors of industry that are major sources of pollution.

- The bulk of the transportation sector work has been carried out by the automobile industry, through cleaner and more efficient engines and improved vehicle design. It is now up to the planners (and the policy makers) to do their part and curb overall highway travel.
Figure 1-2: Growth in Travel, 1950 to 1988. Source: US DOT (1990)
1.2 Traditional Approaches of Dealing with the Problem

1.2.1 Supply Augmentation

The traditional response to congestion is the addition of new highway capacity through construction, and despite the near completion of the Interstate Highway System, some cities adding capacity by extending urban freeways. The general trend, however, discourages new construction, since it is very costly, disruptive to existing traffic, opposed by numerous groups on environmental grounds, and physically constrained. Hence more attention has been given to measures that will increase the effective utilization of the existing system, including urban freeways, arterials and local streets (Meyer et al. [1989]). These measures include, for freeways:

- Ramp metering: Using a traffic signal placed at the end of a ramp the volume of traffic entering the freeway is controlled, in response to the volume of traffic on the ramp or on the freeway. The additional delay due to the ramp signal has been observed to be outweighed by the increased speeds on the freeway.

- Geometric modifications, such as the use of shoulder lanes for travel during peak hours, or the reduction of lane widths to provide additional lanes within the existing
pavement.

- Motorist information systems, such as variable message signs that inform drivers about traffic conditions and (perhaps) alternative routes\(^3\).

For arterials and local streets, these measures include traffic signal improvements or intersection improvements, turn prohibitions, one-way streets, and parking management, to name a few.

1.2.2 Demand Management

One of the drawbacks of supply augmentation techniques is the commonly observed phenomenon that after capacity is increased, congestion levels soon return to those prior to the change. This is usually explained as being the result of latent or pent-up demand, for instance from other modes, such as transit. Although it may be argued that this represents a net gain, since more people are now able to use their preferred choice (highway) at the same levels of congestion, there is nevertheless a general trend away from measures that increase vehicle-miles due to environmental and energy concerns. Another way to reduce congestion is through demand management.

Specific Approaches

In contrast to supply augmentation, demand management seeks to change the demand for travel, both to mitigate currently experienced congestion and/or to avoid future congestion. These approaches include (Meyer et al. [1989]):

- Ridesharing in the form of carpooling, vanpooling, or buspooling. Primary means to encourage ridesharing are employer-based incentives and HOV-lanes and facilities.

- Alternative work-hours such as staggered hours, flex-time, or compressed work weeks. Such programs are employer based and their success depends on factors related to employment and city structure.

- Auto restricted zones, where automobile travel in parts of a region is restricted or controlled in some manner.

- Broad policy measures such as growth management and trip reduction ordinances.

---

\(^3\)Information systems are included in this category because the expected improvements in travel times are achieved using the existing highway network, as if the effective supply had increased. However, since they are used to modify travel behavior, they could also be included among the demand management techniques discussed below.
1.3 Innovative Approaches to the Congestion Problem

- Congestion or peak-period or road pricing, which uses pricing to make driving at the most congested times less attractive, thus encouraging switching to other modes, times or routes. For a substantial period of time it has been perceived to have the greatest potential for modifying demand so that congestion is reduced. Unfortunately, it is also a technique for which public opposition has been strongest.

Basic Principles of Congestion Pricing

The basic theory of congestion pricing is both simple and illuminating. Travel on highways is subject to congestion: in general, the more users, the greater the travel time for each user. The average (private) cost experienced by the user and the (marginal) social cost differ in this case; the difference represents the (unpaid) externality that each user imposes on the other users, and should be charged to the users. The result of this marginal cost pricing is the move to the system optimum.

In general, individual adjustments to congestion pricing schemes include change of route, mode, or departure time, and the decision to travel. In this research, we will primarily employ dynamic models that permit changing the route or the departure time.

1.3 Innovative Approaches to the Congestion Problem

The traditional approaches to combatting congestion reviewed above are not sufficient to solve the congestion problem. Supply augmentation in the form of highway construction is seldom acceptable because of huge costs of construction and environmental concerns. Tested demand management techniques such as ridesharing have contributed on a lesser scale than was hoped (Meyer et al. [1989]).

Advanced Technology

Many in the transportation community have concluded that the solution lies in technology, in particular in harnessing the opportunities offered by advanced communications and computing technologies. Thus, concerted and large-scale efforts were begun in the early 1980’s in Europe and Japan, and more recently in the U.S., to develop comprehensive applications of advanced technologies to improve the efficiency of the highway network that include many capacity-boosting measures. Due to the size of the effort, we will now discuss these initiatives in somewhat more detail.

In the U.S., the term used for these applications is Intelligent Vehicle-Highway Systems (IVHS); in Europe the preferred term is Road Transport Informatics (RTI). Although these programs are broad and diverse, the overall scope is similar, with a difference in emphasis. The PROMETHEUS (Program for European Traffic with Highest Efficiency and Unprece-
mented Safety) program emphasizes on-board vehicle functions, and the DRIVE (European Dedicated Road Infrastructure for Vehicle Safety) program places more emphasis on the transportation systems approach. By 1994, it is expected that almost $1 billion will have been spent on these two projects. In Japan the issue is approached largely on a project basis with the goal being the development of traffic management and traveler information systems. The Road/Automobile Communication System (RACS) and Advanced Mobile Traffic Information and Communication System (AMTICS) are the two main programs in Japan (IVHS America [1992]).

Recent Developments in the U.S.

The U.S. had a somewhat later start into the IVHS field. In 1989 an informal group called "Mobility 2000" published the proceedings of a workshop on IVHS. The group identified the need for research, development, demonstration programs, and implementation strategies. Further work resulted in the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA) authorizing $660 million for IVHS-related research until 1997. The group that currently conducts most of the nation-wide planning in this field, IVHS America, has recently published a guide to development and deployment of IVHS in the U.S. (IVHS America [1992]). Since it contains the latest definitions, we will draw from it with regard to the objectives and main functional areas of IVHS. Thus, the main operational goals include improved safety, reduced congestion, increased and higher-quality mobility, improved environmental quality and energy efficiency, and improved productivity. Institutional objectives include:

- The establishment of a viable and profitable U.S.-based IVHS industry
- The redirection of the transportation profession (particularly relevant with the completion of the Interstate Highway System), expansion of the capabilities of existing organizations and the attraction of new organizations into the transportation field
- The development and demonstration of a new institutional structure for technology development and deployment in the U.S.

The main functional areas of IVHS include:

- Advanced Traffic Management Systems (ATMS): These employ innovative technologies and integrate new and existing traffic management and control systems in order to be responsive to dynamic traffic conditions while servicing all modes of transportation. Key features are subsystem integration and real-time control adjustments.
- Advanced Traveler Information Systems (ATIS) acquire, analyze, communicate and present information to assist travelers in moving from origin to destination. The
systems provide such assistance in a manner that best satisfies the traveler’s needs for safety, efficiency, and comfort. The travel may involve a single mode of transportation, or it may link multiple modes together during various parts of the trip.

- Advanced Vehicle Control Systems (AVCS) combine sensors, computers, and control systems in vehicles and in the infrastructure to warn and assist drivers or to intervene in the driving task. Higher safety levels and lower congestion are envisioned in what could potentially lead to entirely new concepts for surface transportation services. Amongst the main features are sensors that augment human perception, and automated controls that are superior to human reflexes. Examples are warning systems for backup or longitudinal collision. Although some of the most futuristic aspects of IVHS are found in this category, a prototype automated freeway could be demonstrated as early as 1997, based on ISTE A.

IVHS and Congestion Pricing

Two things should be noted about IVHS. First, IVHS techniques are generally not new in principle. They represent improved methods of carrying out existing functions of the transportation control systems. Radio information on traffic conditions, for example, has been available for decades, and modern ATIS simply represents a systematic and technology-intensive way of carrying out the same task. Second, IVHS focuses on supply augmentation techniques, and is therefore likely to suffer from the induced demand problem in the same way as other capacity-boosting techniques, prompting concerns that the promised benefits will simply not materialize.

In our view, IVHS will prove most fruitful when combined with demand management strategies such as congestion pricing. In simplified terms, IVHS can extract more capacity from an existing network, while congestion pricing will both help to improve current utilization and deter excessive induced demand. Moreover, IVHS is likely to play an important role in the implementation of congestion pricing, making the latter one type of “Advanced Demand Management Technique”. Currently available technology permits pricing that is unobtrusive, flexible, guarantees privacy, and can be carried out at highway speeds without generating delays. Such electronic toll collection systems are in place in numerous states and performing well.

*These systems are used to collect ordinary time-invariant tolls.*

27
1.4 Research Objectives

Some of the advanced congestion alleviation techniques discussed in section 1.3, such as ATIS and ATMS, appear particularly useful for reducing incident congestion. We believe that the policy with the greatest potential to reduce recurrent congestion is congestion pricing and that combining congestion pricing with IVHS tools of the type mentioned above represents a comprehensive approach to combat urban congestion. To investigate the validity of these beliefs, it is proposed that the effects on congestion of a combined congestion pricing and traveler information system is determined, for several reasons. First, the results may be indicative of the general trend of jointly implementing pricing and control schemes. Second, the concomitant introduction of the information system can be perceived as providing a new product to accompany the new charges, thus avoiding the situation where commuters are charged with "nothing" provided in return, which occurs if pricing is introduced alone. The information system is a "palpable" benefit and can complement or substitute for the redistribution of toll revenues. Finally, the presence of congestion pricing might result in lower levels of recurrent congestion. Intuitively, the re-routing that occurs in the presence of information systems may be more likely to result in system gains when there is more excess capacity on routes unaffected by incidents; there may also by synergistic effects in the combined system.

Thus the objectives of this research are:

1. To review and develop evaluation criteria related to the practical implementation of congestion pricing schemes, centered around an analysis of current trends and an examination of the theoretical literature, in order to identify potentially successful schemes. In this research, implementability issues related to congestion pricing that will receive particular attention are the temporal and the spatial toll structure. In addition, implementability is discussed for the case of information systems and for combined pricing and information systems.

2. To identify design issues that are particular to the selected combined pricing and information systems.

3. To evaluate some of these issues in a quantitative setting, using simulation together with an equilibrium model. We intend to explore the modification of both the information system and the pricing scheme to overcome potential implementation problems. We also attempt to determine whether the benefits of a combined system are greater than the sum of the benefits obtained by applying pricing and information separately.
1.5 Thesis Outline

Chapter 2 reviews the literature on congestion pricing, traveler information systems, and combined pricing and information systems.

Chapter 3 begins with a characterization of the different congestion pricing and traveler information systems. Following an explanation of the role we think pricing and information should play in combating the various forms of congestion is the presentation of an evaluation framework for assessing various aspects of congestion pricing. This is used to select a particular scheme for further analysis. It is also extended to select the appropriate information system to be combined with the specified congestion pricing scheme.

In Chapter 4, the reasoning behind our choice of modeling approach is explained, and a set of integrated design issues described. This is followed by a complete description of the model used to perform the quantitative analysis.

Chapter 5 presents the results of the analysis, including the effect of modifying the information system. The superadditivity of pricing and information policies is explored, followed by evaluating the impact of imposing additional implementability constraints on the system through a modification of the pricing schemes. In particular the effects of imposing identical toll structures on the different routes, and the availability of an untolled alternative, are determined.

Conclusions and directions for future research comprise Chapter 6.
Chapter 2

Previous Research

This chapter reviews the literature on congestion pricing, traveler information systems, and combined pricing and information systems. By necessity, it is somewhat selective, since the literature is vast, particularly that on congestion pricing. Emphasis is therefore given to reviewing recent research that is of direct relevance to this thesis.

2.1 Congestion Pricing

In this section, both the theoretical literature and some practical implementations of congestion pricing are reviewed.

2.1.1 Theoretical Literature

Economic theory has provided the basic principles behind congestion pricing\(^1\). The first concept is that of externalities. One of the reasons why a competitive price system may fail to achieve economic efficiency is the presence of externalities, which are interactions among

\(^1\)One point to be made relates to the possibly divergence between the meaning that the term “congestion” carries to economist and transport planners. The economists view of congestion is expressed as the increase in cost of use of a facility due to an increase in the number of its users. If it is possible to charge users for the congestion externalities they impose on others, a system optimum will be reached, at which point congestion is eliminated.

The transportation engineer’s or planner’s view of congestion is more empirical. For example in the U.S., congestion on freeways is described using the concept of level of service, which is based on the density of traffic on a segment (i.e., total number of vehicles) which in turn is related to driver’s speeds and their ability to maneuver within the traffic stream (Transportation Research Board [1993]). Six levels of service from A to F are defined for the various ranges of average speeds and densities, and hence there is no clear-cut point at which congestion is said to begin.

The important consequence is that it is possible that eliminating congestion in the economist’s sense, which is the goal of congestion pricing, may result in levels of congestion that are still high from the perspective of transportation professionals (and the public). Frequent claims by transportation professional who look for optimal tolls as those that will eliminate all congestion suggests there is some confusion in regard to this point.
firms or individuals that are not adequately reflected in market prices (Nicholson [1993]). Credit goes to Pigou (1920) for pointing out one possible remedy, that of taxation, and for demonstrating its application to congestion externalities through an example consisting of two roads. The common aspect of all congestion pricing models is the search for a toll that achieves a social optimum based on the appropriate definition of user cost.

The preferred approach used by economists has generally been to assume the existence of a demand curve. On the other hand, research carried out by transportation planners is driven by the desire to model the component choice processes that will constitute the response to pricing, and to solve the problem for general networks. Most of the literature consists of the economists approach, sometimes using a simplified network representation. We will look at both the demand curve and the component choice process approaches below. The related literature of peak-period pricing will also be discussed in terms of its relevance to congestion pricing.

Using a Demand Curve as Aggregate Response

The greater part of the literature utilizes a demand curve to derive the optimal toll. Using that approach, a demand curve is assumed to exist, capturing the demand on the facility as a function of price during the peak-period. Presumably, the demand curve implicitly captures the interaction between traveling during the priced period, and adjusting to that price through canceling trips, switching modes, or even traveling at other times. The optimal fee is given by the difference between the point of intersection of the demand curve with the marginal cost curve, and the average cost curve at that point, as shown in Figure 2-1.

This approach has been used by Walters (1961), Vickrey (1963), Johnson (1964), Keeler and Small (1977), Harrison (1986), Braid (1989), and Kraus (1989), among many others. This model appears too vague on the issue of predicting actual responses to the pricing scheme to be effective for assessing proposed implementations. The approach has been widely used because of its relative simplicity compared to the more complex models described below. This includes its application to policy studies, where the estimates of the relationship between speed and costs are used to make the theory operational (Walters [1987]).

Whether congestion pricing is equitable has also been analyzed using these models. The predominant practice has been to show, for the particular model being analyzed, that some people are "made worse off", that congestion pricing is a regressive or progressive tax, or that everyone benefits when pricing is implemented.

Thus Layard (1977) analyses the Walters (1961) model, allowing for two classes of commuters, poor and rich. He concludes that, if the disposal of revenues is ignored, the congestion tax may be regressive as it discourages journeys with low travel time values and
probably encourages those with high time values. Richardson (1974) argues that congestion taxes are regressive in their effects on car owners but ambiguous in their impacts on all road users (including road-based transit), and that equity questions as a whole cannot be resolved. Glazer (1981) argues that consumer surplus is greater in the no-toll case than if optimal tolls are levied, for the simplified model of a taxicab market. However, the reduced volume that these models assume to occur as a result of pricing, as shown in Figure 2-1, can be the result of a variety of adjustments, as pointed out above. Therefore the conclusions these models draw with regard to equity appear of limited value.

Directly Modeling the Response to Pricing

Route Choice Another group of modeling approaches explicitly considers the component choice processes which actually form the demand curve. One of these components is the route choice, and the optimal toll in this case is defined as that which brings about a social optimum with regard to route choices (ignoring other possible adjustments). We begin with an illustrative example.

Suppose we have a simple network, depicted in Figure 2-2, with a single origin-destination pair, and a fixed total demand $N$. Assume that travelers make their route choice decisions so that user equilibrium results: no traveler can improve their travel time by unilaterally switching routes. The private cost experienced on each route as a function of the number of users on it is given by the lower curves $A1$ and $A2$ in Figure 2-3 respectively, and the corresponding marginal cost of travel is given by the curves $M1$ and $M2$. Prior to
Figure 2-2: A two-route network with inelastic total demand

pricing, the system reaches equilibrium at the flows corresponding to the intersection of the average cost curves, namely point \( UE \). If we charge routes 1 and 2 the difference between the system optimum point \( SO \) and the value of the appropriate average cost curve at that point, the congestion externality would be internalized, and at the new user equilibrium, marginal costs would be equal across routes and hence correspond to the system optimum.

Dafermos and Sparrow (1971) outlined the problem for general networks, and solved it on small examples. They derived optimal route-based tolls which form the exact analogue to the simpler cases (with single link networks) since they permit charging each commuter the exact marginal cost of his trip. An important observation is that (the more practical) link-based tolls that do not distinguish between users' origins or destinations can not, in general, reproduce the route-based tolls necessary for total cost minimization, except for simple networks, even when we can charge tolls on all links\(^2\). Thus charging users the correct toll on a general network through link-based tolls that do not differentiate between travelers appears infeasible. In practice one must resort to the suboptimal task of charging a toll on only some links of the network, not differentiating between commuters with different origins and destinations. However, commuters may only have a handful of routes from which they choose in practice, or are willing to use. It is not clear whether under those circumstances the restrictions on link-based tolls would still apply.

Boyce et al. [1990] discussed the applicability of bilevel programming techniques to solve the problem of charging the best possible tolls on an intercity road network during predefined time periods and on specific zone-crossing links, with reference to a proposed

\(^2\)Their conclusion is that "the link-toll formulation is solvable in general (i.e. disregarding improbable redundancies) if and only if the number of paths is less than or equal to the number of links plus the number of OD pairs", a condition which is unlikely to be satisfied in practice as network size grows.
Figure 2.3: The theory: Charging the difference between social and private cost results in system optimum
Dutch scheme. They suggest a number of upper level program objective functions to be minimized, in addition to the total auto travel time, or total auto and transit travel time, over the whole network. The lower level program consists of user equilibrium formulations of route choice, departure time period choice, and mode choice\(^3\). The model is hence similar to the network design problem, which is very complex even in the static setting. Thus, for designing tolls on an urban network, one could use a similar approach as that suggested by Boyce et al., modifying it in order to model departure time choices, and minimize an appropriate objective function. Without simplifying assumptions it becomes not solvable currently, due to its complexity and because no workable approaches to dynamic network equilibrium problems have yet been developed.

**Departure Time Choice** By introducing the modeling of the *departure time* choice dimension, a seminal contribution to the modeling of time-variant pricing was made by Vickrey (1969), who derived optimal time-varying tolls. In his model, there exists an equilibrium among travelers as a result of the trade-off between arriving early or late at work and incurring little travel time, and arriving close to the desired arrival time but incurring greater travel time. Previously, it was usually assumed that those travelers who did not want to pay the toll would switch to other modes, or at best to other "periods". Vickrey's model suggested that they could switch departure times in response to a time-varying toll. The conceptual development lay in the assumption that the commuter can trade off travel time against *schedule delay*, which is the difference between actual and desired arrival time at work. This provided a tractable way to model user response to tolls, even for traffic that is inelastic in regard to mode choice. It spawned a great deal of work that incorporated its main assumptions. An example is Henderson (1974), who replaced the pure bottleneck congestion assumption with a traffic flow model based on speed-density curves. In that context it should be noted that the dynamic user equilibrium problem for general networks has not been satisfactorily solved, even for the deterministic case, nor is there agreement on precisely how the problem should be approached (Watling and van Vuren [1993]).

Of particular interest is a deterministic model based on the concept introduced by Vickrey that has been developed by Arnott et al. (1990a, 1990b). It examines the effect of time-varying, uniform and step-tolls on user benefits for single and parallel route networks. We have used it to generate default or average behavioral departure time and route choice patterns that capture the effect of the congestion pricing schemes. The model and our use of it will be described in more detail in Chapter 4.

\(^3\)Note that the trip is assumed to occur in a single time period, and thus departure time adjustments to pricing are not modeled.
2.1 Congestion Pricing

Peak-Period Pricing

Related to the literature on congestion pricing for transportation is work on the theory of peak-period pricing. Among the classical references are Boiteux (1960), Steiner (1957), Williamson (1966) and Littlechild (1970). In public utilities including electricity, gas, water, and telephone services, the theory of peak-pricing has been applied to modify demand patterns. The theory is motivated by the attributes of public utilities’ products, which are economically non-storable and the demand for which varies over time. The non-uniform utilization that may result is discouraged by implementing a peak-load pricing policy which reduces consumption in peak periods and encourages off-peak consumption (Crew and Kleindorfer [1986]).

The standard model assumes no congestion, i.e. the quality of the product is not affected by the level of demand. The only objective is to reduce peak-period demand and thus the capital equipment needed to provide that demand. This model has evolved to include congestion costs, i.e. the quality of the product consumed decreases as the demand for it increases (e.g. Cohen [1987]). It divides the time frame under consideration into several periods. A separate analysis is carried out in price-quantity space for each of these periods. The quantity is the number of facility users, and the price is the user’s private cost, which includes costs due both to tolls and to congestion. Again, marginal social cost is given by the increase in social cost due to an additional user in that period. The demand curve depends on both the time period and the equilibrium prices for the other time periods, due to inter-temporal substitution. The social optimum is at the intersection of the demand and marginal social cost curves.

Unfortunately, this model is not appropriate for highway congestion pricing. Some of its ambiguities, when compared to actual highway congestion phenomena are: First, it is not clear what constitutes the appropriate number of users in a time period. Second, the private cost in a time period is also a function of the number of users in other time periods (assuming the total period under consideration is one in which congestion does not completely abate). Third, the addition of a driver in a time period also increases delay in other time periods (Arnott et al. [1993]).

In spite of these limitations, the peak-period pricing literature is potentially relevant to the transportation application. The current emphasis in congestion pricing is on time-varying pricing that can be viewed as replicating the objectives of peak-period pricing - namely to flatten the demand curve, though for different reasons. The practical experience of applying time-varying pricing in these industries may prove to be of use when implementing congestion pricing.
A Note on Terminology

There are several terms that are usually used to describe the time-varying pricing which forms the subject of this research. We will attempt to restrict ourselves to the use of the term *congestion pricing* in this thesis, and we will be referring to a *time-varying pricing mechanism designed to induce behavioral adjustments that will reduce some measure of congestion*. However, this term is not without ambiguities - congestion also can be reduced through time-invariant pricing. *Road pricing* is another frequently used term, but it evokes the image of time-invariant tolls designed to generate revenue for the maintenance and operation of highways (although time-variant tolls also generate revenue - thus it is more a matter of intent and focus). *Peak-period pricing* is reminiscent of the same-named theory, which is unsuitable for direct application to highway congestion, as discussed above. *Marginal-cost pricing* is also not quite appropriate, because we are only pricing for one externality and ignoring, for example, pollution and road deterioration⁴, and because in an actual application it is unlikely that anyone can be charged precise marginal costs. It would probably be best to use the term *pricing for congestion reduction*, as defined above, but this is too cumbersome.

2.1.2 Practical Applications of Congestion Pricing

Following Pigou’s (1920) lead in calling for marginal-cost pricing for highways, increasingly concrete proposals for its actual use were made in the 50’s and early 60’s by Walters (1954) and Vickrey (1963). Testimony in favor of pricing was made to a Congress subcommittee by Vickrey (1959). The most important step towards implementation was made in the U.K. in 1962, when the government appointed a committee to study the economic and technical feasibility or road pricing. This resulted in the well-known “Smeed” report (Ministry of Transport [1964]), a fairly detailed account of important aspects of designing a workable policy, which included the suggestion to use automatic meters. Although substantial net benefits were predicted for the introduction of pricing, the policy was never adopted.

Things were quiet for a while, but in the 1970’s interest resurfaced. The City of London considered charging vehicles for their presence in a central zone as a proxy for their contribution to congestion, but did not move beyond the proposal stage. In the U.S., several cities were offered assistance to implement road pricing demonstrations through a 1977 solicitation (Higgins [1986]). Three expressed interest (Berkeley, California; Honolulu, Hawaii; Madison, Wisconsin); however, preliminary studies ended without request for further study and possible demonstration of road pricing. The reasons for rejection were primarily related

---

⁴See Newberry (1989) for a discussion of marginal cost pricing applied to externalities other than congestion.
to the economic, political and administrative obstacles of implementation, which included concerns in regard to excessive and regressive taxation, and negative impacts on business. Similar reluctance to implement even experimental projects has been the trend through much of the world, including, among others, Japan, Thailand, and Malaysia (Armstrong-Wright [1986]). The only success story during that decade was the Singapore congestion pricing scheme, begun in 1975 and still in place today. We will discuss this scheme in more detail below.

The last decade has seen what can be viewed as a third wave of interest in congestion pricing. It is motivated by congestion that appears only to have worsened, increasingly visible impacts of highway traffic on the environment, and the availability of relatively inexpensive advanced technology. The first project was the Hong Kong pilot project, initiated in 1983 but not fully implemented. This was followed by plans to introduce congestion pricing (once again) in Greater London. In Sweden and Norway, cordon rings around major cities have been installed, and, while they are currently collecting time-invariant tolls, are poised to adopt peak-pricing toll structures (Hau [1992]). The Netherlands is in the process of developing a pricing scheme, as is the city of Cambridge in the U.K. In the U.S., ISTEA has authorized funding for up to five congestion pricing pilot projects, and already one has been approved, the San Francisco Bay Bridge.

We review some of the more interesting proposed or actual schemes in more detail below. In addition to the Singapore project, we will look at the Hong Kong experiment, the proposed schemes of the Netherlands and Cambridge, and the projects underway in the U.S.

The Singapore Area Licensing Scheme

The best known existing congestion pricing scheme is the Singapore program, begun in 1975, where initially all drivers entering the CBD between 7:30 AM and 9:30 AM were charged a fixed toll of $1.30/day [Holland and Watson (1978)]. Due to congestion occurring after 9:30 AM, the toll period was subsequently extended to 10:15 AM and the toll was also increased to $5.00/day. Private cars with more than 4 occupants, goods vehicles, and buses are exempted from the fee. The program is viewed as a great success, enforcement is effective, and most important, acceptance by the public and the business community has ensured its continuation.

Initial goals were to reduce traffic by about 25-30%. It is not clear that this was an optimal figure, assuming you can meaningfully define "optimal" in such a context. Actual reductions were of the order of 40-45%.

The system has evolved over the years. In 1989, major changes included charging during the afternoon peak period from 4:30 - 6:30 p.m. (initially 7:00 p.m. but changed to
accommodate businesses). Carpools were no longer exempted from the toll, and the daily license fee was lowered from S$5 to S$3. Motorcycles must pay S$ 1.00 per day.

The latest development is the decision to implement electronic road pricing that will replace the windshield-sticker based system currently in use. Field demonstrations for various systems are to take place in May 1994. Implementation of the selected system is scheduled for 1997.

The Singapore project continues to serve as an example to the rest of the world on the potential of congestion pricing. Recent studies, (Wilson [1989], IcCarty and Tay [1993]) however, have questioned the achievements of the scheme. Primarily, they cite absence of congestion during the toll period as an indication that the toll levels are set too high. As Holland and Watson (1978) remark, it is not easy to assess the benefits due to the pricing scheme, since only months before the introduction in 1975 a stringent program to discourage automobile ownership was also implemented.

**The Hong Kong Electronic Road Pricing Study**

A scheme that was implemented successfully as a pilot project (1983-85) is the Hong Kong program (see Dawson and Brown [1985], Dawson and Catling [1986], Harrison [1986], Dawson [1986], Harrison et al. [1986], Catling and Roth [1987]). Based on electronic road pricing like the proposed Dutch program, each vehicle in the program was equipped with an electronic number plate which was interrogated using a multi-lane loop array buried beneath the road at specific toll stations. There were 2500 vehicles in the project, which were tolled at 18 toll stations, the toll depending on the time of day and the location, and several zones, with a peak, intra-peak, and off-peak charge applied for crossing the zones. Drivers were billed for their road use on a monthly basis. The government of Hong Kong proposed adoption of a full-scale electronic pricing scheme in 1985.

The proposal was overwhelmingly rejected due to public resentment because of concerns over privacy and excessive taxation, and because of business concerns. Ongoing congestion problems (Hong Kong has the highest vehicle concentration in the world) have led to the reconsideration of electronic road pricing in Hong Kong (Hau [1990]).

**The Netherlands project**

Until recently the most significant program of congestion pricing under consideration in Europe was the Dutch Ministry of Transport scheme [see Stelhorst and Zandbergen (1990)]. Their stated goals were to reduce the long-run growth of car use, to reduce waiting time costs, and to eliminate congestion. The proposed system included equipping vehicles with prepaid smartcards in an on-board communication unit. The plan was to divide the country into cells and to debit the cards automatically, without the vehicle stopping at a toll booth.
The toll would vary with the time of day and the cell, with specific reduction licenses for some, such as trucks or disabled persons. Due to political difficulties the system's proposed implementation, scheduled for 1992, has been rejected. It has survived in the form of a smaller version planned to be introduced by 1997 (Hau [1992]).

The Cambridge Road Metering Project

In 1990, the Chartered Institute of Transport published a report on congestion and road use charges entitled "Paying for Progress". One of its conclusions was that, to maximize benefits, road pricing should be introduced in London, but it was desirable to gain experience with a scheme in a smaller city. Cambridge, a city which has major congestion and traffic growth problems\(^5\) has proposed to do exactly that. As part of a comprehensive package of measures that were initially advanced in 1989 in the Cambridge Transport Strategy and include transit improvements, extension of park and ride sites, and addition of road capacity, it was decided in October 1990 to develop plans for "congestion metering". A synonym for a form of continuous pricing, the scheme is designed to work as follows:

Vehicles are equipped with meters which automatically begin recording time when the vehicle enters the tolled zone. For every 0.5 km traveled in the area, the meter checks whether the time taken to travel that distance exceeds the specified threshold. If it does, the unit congestion charge is levied to the vehicle's smart card; otherwise, no charges are made. The process continues in this fashion until the vehicle exits the zone, or parks (turns off the engine). Thus vehicles are charged approximately for the distance traveled in congested traffic. The system is expected to begin experimental implementation in 1995 (Hills and Blythe [1992]).

Because the system charges a fixed amount for traffic levels above certain points, pricing is clearly very crude, since the typical congestion curve is effectively replaced with a straight line above the appropriate flow level. To see this clearly, assume we are only looking at the possible charges arising for the travel between \(n\) beacons. Now, a fixed charge is applied if either a certain number of stops is made, or the time exceeds a certain length; the first criterion is more arbitrary than the second, but let us assume it corresponds to a single level of congestion. When either of the two levels of congestion corresponding to the two criteria (which should be close to each other) are exceeded respectively, a toll \(\tau\) is charged.

\(^{5}\)During the period from 1980-1990, traffic grew by 50% on the main radial routes into Cambridge, and traffic levels are expected to increase by a further 40% by the year 2000. These predictions are based on the Department of Transport's national GDP forecasts, and make allowance for the above average population and employment growth as well as the imbalance between the number of jobs and dwellings within the city (Hughes and Ison [1992]).
Projects in the U.S.

Interest in experimenting with congestion pricing in the U.S. is strong again. San Francisco is the first city to qualify for federal assistance from the ISTEA budget for the nation's first congestion pricing project. A system of variable tolls for drivers crossing the San Francisco-Oakland Bay Bridge is to be tested. The project has begun in September 1993, with actual implementation of time-varying tolls scheduled for 1995. Stated objectives are to reduce congestion during peak hours and increase demand during off-peak hours. The project will examine alternatives such as non-carpool vehicles paying higher tolls during peak morning and evening commute hours (Urban Transportation Monitor [1993]).

Orange County in Southern California has plans to construct five new toll roads which will use automatic vehicle identification, automatic toll collection equipment, and changeable message signs to guide traffic. Two of those are proposed by private firms and will apply congestion pricing both to reduce peak-period demand and to increase revenue. The effort is facilitated by the recent passage of a bill that legalizes private highways in California (Fielding [1993]).

2.2 Advanced Traveler Information Systems

Static and dynamic network equilibrium models generally assume that commuters possess perfect information on the state of the network. That assumption is obviously a great simplification of reality; in practice, travelers have imperfect information on travel times on alternate routes, location and severity of traffic incidents etc., a situation which is aggravated by the high stochasticity of an urban transportation network. The essential aim of traveler information systems is to reduce this "information gap" and to improve some system performance measures. Of course, the practice of obtaining information on traffic conditions for the purpose of making better travel decisions is not new with regard to such sources as word-of-mouth and radio messages.

However, Advanced Traveler Information Systems (ATIS) attempt to provide information of this nature in a more systematic and potentially beneficial fashion. These systems provide information to travelers in their vehicles, in their homes or at their workplaces. The information might include the location of incidents, weather problems, road conditions, optimal routings, lane restrictions, and in-vehicle signing (IVHS America [1992]). In contrast to congestion pricing, ATIS has received much more in the way of practical application, but less theoretical study. We therefore begin with a review of some existing practical schemes,

---

6ISTEA provides for the funding of up to five pilot projects.

2.2.1 Practical Applications

Pilot implementations of ATIS are now in place in much of the industrialized world.

Japan One of the first was the CACS project, which was begun in Japan in 1973 and completed in 1979, in which vehicles were provided with shortest travel time paths based on predicted route costs. Savings in travel time to equipped vehicles were estimated to range from 9% to 14%. In that country the two main projects today are the Road/Automobile Communication System (RACS) and the Advanced Mobile Traffic Information and Communication System (AMTICS). The former is oriented towards expressways and the latter towards surface streets. RACS is designed for exchange of information between the control center and individual automobiles. Currently both systems are designed to provide real-time traffic information without suggesting particular routes. However, plans have been made to provide shortest-route directives based on that information, and to integrate the two systems under the VICS name.

Europe ALL-SCOUT and AUTOGUIDE are two projects that are part of the DRIVE project EUREKA. ALL-SCOUT is a route guidance system developed in Germany which uses infrared transmitters and receivers to transfer navigation information between roadside beacons and on-board displays in equipped vehicles. An advanced version of this system is being tested in Berlin under the LISB acronym, with beacons installed at 250 selected intersections and at other locations which transmit location, map, and shortest path routing trees to 700 equipped vehicles as they pass the beacons. In-vehicle units suggest the shortest route through simple graphics together with audio messages based on the driver's indicated destination. The route guidance is based on predicted journey times downstream, which are obtained by updating previous or average times with the information flowing from the vehicles to the control center through the beacons. EURO-SCOUT is the name for the system that has grown out of the ALL-SCOUT project, including public transport and traffic management elements. This concept is being further tested in pilot projects in other German cities. AUTOGUIDE is a route guidance system developed in England based on the ALL-SCOUT technology, and after a small pilot project is completed, a larger one is planned to be completed in 1993, with 1000 equipped vehicles.

U.S. The main projects are: California's Pathfinder in the Smart Corridor of Los Angeles, which is a small-scale project that provides information on traffic conditions with no route directives. Travtek in Orlando, Florida is currently being tested, providing navigation help
through graphic displays and audio messages and by providing information on location of traffic incidents and congestion. Route guidance is based on current traffic conditions. A larger project is planned in the Chicago area in the form of the ADVANCE project, where about 4000 vehicles are to be equipped with in-vehicle units, and traffic-dependent guidance is expected to be provided to drivers.

2.2.2 Theoretical Work

Mahmassani and Herman (1990) have developed a useful hierarchy of approaches used for studying large-scale interactive dynamic systems involving individual decision-making, such as those which arise with advanced traveler information systems. These are

- Analytic models of idealized situations
- Simulation models, under assumed micro-level rules, parameters, and hypothesized relations.
- "Laboratory" experiments: controlled experiments, limited number of experimental factors, manageable number of participants, and simulation-supported interactions.
- Field surveys.
- Field experiments: some aspects of the system is changed and user response is monitored.

In general following this hierarchy implies increasing cost and realism or validity, decreasing control over the system and ability to monitor events, and increasing difficulty of obtaining simple fundamental insights about system behavior.

Analytic models of Idealized Situations  An example of the first category is Arnott et al. (1989), who analyzed traveler information systems using closed-form models of equilibrium situations, and concluded that benefits do not always accrue. This line of work is generally too abstract to permit useful generalizations.

Simulation Models  A great deal of research on information systems has been carried out using simulation models. For example, Tsuji et al. (1983) developed a stochastic approach to estimate the effectiveness of route guidance systems. The benefits are quantified in terms of various measures of effectiveness applied to the guided vehicles. These measures include the probabilities of the guided vehicles arriving at their destinations earlier than the unguided vehicles, and the travel time reduction ratio for the guided vehicles. The simplifying assumption is made that the flow of guided vehicles does not affect the remaining
traffic flow, which is justifiable when the number of guided vehicles is small. Based on these estimates, predictions of the monetary benefits to be obtained by implementing information systems are made.

Using a somewhat different approach, Koutsopoulos and Lotan (1990) examined the impact of information systems on recurring congestion levels by assuming that the variance of the perceptions of link travel times decreased as a result of supplying information. A stochastic traffic assignment model was used in a static setting, and applied to a small network. Average reductions in total system cost of 4.4% were reported.

Mahmassani and Jayakrishnan (1991) use simulation to evaluate system performance of a traffic corridor when various assumptions on user response to information are made. The impact of varying the fraction of guided vehicles is also explored. It is shown that myopic local actions by individual drivers may actually result in worse outcomes for them, as well as systemwide losses.

The effect of various design parameters, such as the spatial and temporal update frequency, the guidance rate, and the methodology for information provision, on the performance of an information system in a simple network was explored by Kaysi (1992). Important findings of that research were that information must be predictive, and that the expected user response to information must be taken into account.

Laboratory Experiments "Laboratory" experiments that use human participants who respond to various travel scenarios generated by traffic simulators are becoming a popular tool for exploring traveler response to various information system configurations and characteristics. They are being developed at several research centers around the world. An important aspect of this work attempts to evaluate the response of travelers to imperfect, or occasionally unreliable, information. Since perfect systems are considered a thing of the future, this is a most important design issue. Simulators are being developed at several universities in the US, Europe, and Japan. An example of recent work is that of Mahmassani and Herman (1990).

Field Surveys Revealed preferences can be obtained from field surveys. A major effort in this regard has recently been undertaken by a group of researchers at the University of Washington. A survey was designed that focused on motorist behavior and decision processes, particularly as they relate to the design and delivery of motorist information. Responses on 62 variables were obtained from 3893 Seattle commuters in September 1988. Amongst the main findings of an early study (Spyridakis et al. [1991]) of the data are: Motorists are more likely to change their routes from work than from home, to divert to known routes sooner than to unknown ones, and to be influenced by traffic information, congestion, and time of day.
Subsequent work (Conquest et al. [1993]) used cluster analysis to identify commute groups with similar patterns of responses to the influence of traffic information. The segments were found to be: (i) route changers, who are willing to change route both en-route and before leaving; (ii) non-changers, unwilling to change departure time, route, or mode; (iii) route and time changers, willing to change route and departure time; and (iv) pre-trip changers, who are willing to change departure time, route, or mode before departure but unwilling to change en route. The three groups of changers were found to make up about 75% of the sample, which suggests that information systems may have a substantial constituency. The study’s results, together with follow-up interviews, have led to the development of design criteria for an ATIS in the Seattle area by Haselkorn et al. (1990).

Further Work

Ben-Akiva et al. (1991) present a theoretical dynamic network modeling framework that incorporates behavioral models of drivers’ route and departure time choices and their day-to-day adjustment processes, and combines it with explicit models of drivers’ information acquisition and integration. The proposed models have not been estimated. In addition the authors point out a number of adverse impacts of information which occur if information is not supplied carefully. These impacts include overreaction, which occurs if too many drivers respond to information on traffic conditions, such as excessive switching to a route in the case of an accident; and oversaturation, a man-machine problem which results when a driver is unable to process the information supplied, which can result in worse off travelers even when the information supplied is accurate.

2.3 Combined Congestion Pricing / Traveler Information Systems

Evaluating the effects of jointly implementing congestion pricing and traveler information systems has been informally discussed by Goodwin (1991), who argued on intuitive grounds that information systems would perform better in the presence of congestion pricing due to lower congestion levels and hence more excess capacity on routes for rerouting. Similarly, Van Vuren and Smart (1990) hope that combining information and pricing would result in synergistic effects, with greater improvements than if each was implemented in isolation. A highly stylized analysis by de Palma (1992) examines the benefits of joint implementation of pricing and information in a game-theoretic setting which, due to its very abstract nature, is of limited value in helping to understand the interaction of the two policies in realistic situations.

In the only other published quantitative work on the topic, de Palma and Lindsey (1992)
have developed analytical models for evaluating the benefits of combined route guidance and road pricing systems. In particular, the evaluation examines whether the gains due to joint pricing and information are greater than the sum of the gains due to the policies applied alone ("superadditive"), less ("subadditive"), or independent. They argue that \textit{a priori} it is not clear which of the two cases to expect. Subadditivity can be expected on account of the principle of diminishing marginal returns. Superadditivity, on the other hand, may be expected for the following reasons: Theoretical studies have shown that information has the potential to exacerbate congestion, and the benefits of "good" information can attract pent-up demand from other modes, annulling the gains. This can be explained by the theory of the second best, since unpriced congestion reflects a market failure in the presence of which a policy such as information provision, which may be beneficial in a first-best world, may no longer yield improvements. Hence, to the extent that tolls guide the user equilibrium towards the system optimum, information is more beneficial with pricing than without.

Three cases are considered: commuters have to decide (i) whether or not to travel, (ii) which route to use, and, (iii) which route to use and when to leave home. Information can be perfect, imperfect, or zero, and relates to the accuracy with which the actual daily capacity is predicted. Independence of benefits due to the policies is found for the first two cases. For the last and most relevant case, superadditivity is found to hold, if the pricing is applied in a state-dependent manner that varies from day to day in accordance with the changes in capacity and also varies over the course of the day; otherwise, for the cases of tolls that are uniform or state-dependent but time-independent, gains are found to be independent. In order to avoid intractability their model assumes:

- Information is supplied before any travel begins.

- That the capacity is constant on a given day, i.e. incidents last throughout the peak-period.

The first assumption is dictated by the network structure, which lacks intermediate nodes between the origin and destination. En-route information systems can be modeled in more complex networks with intermediate nodes between the origin and destination. The assumption of capacities that are constant during a day is made for analytical tractability. It is an unrealistic assumption, since the capacity of a network changes throughout a day in response to accidents, breakdowns, weather conditions, and other factors, and it is this aspect of network performance which to a large degree motivates much of the work on IVHS technologies, such as ATIS and ATMS.

Another assumption which is highly problematic is the existence of equilibrium under full or imperfect information. Since the network characteristics change from day to day, it appears impossible that commuters can to select their departure times or routes in response
to these changes in such a way as to produce a user equilibrium.

Further work is therefore needed in the modeling of combined pricing and information systems. This research attempts to rectify some of the shortcomings of the work we reviewed on this topic. Before presenting the details of the modeling approach used, we need to identify which combination of pricing and information systems to select from the numerous possible choices. This forms the topic of the next chapter.
Chapter 3

Identifying Candidate Combined Systems

This chapter has three main objectives. The first is to characterize the different congestion pricing schemes and information systems so that the set of possible combined systems that need to be considered can be identified. The second is to review important design criteria in congestion pricing and information systems, both for their own sake, and as a tool for achieving the third objective, which is to select a combination of pricing and information systems to be analyzed quantitatively in the subsequent chapters.

Our emphasis is skewed somewhat towards congestion pricing. The reason is that the debate continues on what really forms an effective, equitable, and user-friendly pricing system, with widely varying views. Even the objectives are sometimes not agreed upon; some view (congestion) pricing as a means of reducing the role of the automobile in society, while others see it in the more limited role of a traffic management tool. In particular, the role of the main design aspects of interest to us, i.e. the role of the temporal and spatial toll structure, is often viewed in radically different ways, giving rise to systems that differ significantly in many ways.

On the other hand, the work in information systems is more streamlined. Currently it is far more extensive than the work on pricing, as seen in the numerous projects in development or implementation. However there seems to be a consensus with regard to the basic objective. It appears the primarily focus is on "getting right" the extremely complex task of modeling behavioral responses to information and of producing accurate information, although differences arise on how best to achieve this. Our recommendations in regard to the choice of information system to be examined in this research are therefore reflective of this situation.

49
Chapter 3  Identifying Candidate Combined Systems

3.1  Characterization of Congestion Pricing and Traveler Information Systems

It should be evident from the literature review in the previous chapter that there are numerous ways to implement both congestion pricing and traveler information systems. It is clearly not possible, and, as we hope to show, not desirable, to study them all, since some of the systems that have been suggested or proposed suffer from serious drawbacks whose identification will obviate deeper analysis. We therefore begin by characterizing the set of possible pricing and information schemes so that this set can be reduced in a systematic fashion.

3.1.1  Characterization of Congestion Pricing Schemes

When designing a congestion pricing scheme on a general network, many variables come into play. As shown in Figure 3-1, we need to decide where, when, whom, how much, and how, to toll. That is, locations for toll “stations”, the temporal toll structure, the applicability of tolls to various groups of travelers such as HOVs or taxis, the toll levels, and the technology for toll collection and methods of enforcement are important elements of a pricing policy.

Not all of these factors are examined in this research. We assume that we have advanced technologies for toll collection available, possibly with a known cost of collection, but do not go into technological issues in detail here. Similarly, which classes of vehicles may be exempt from tolls under which conditions is left unaddressed, as is the issue of enforcement strategies. Our focus here is on the temporal and spatial toll structures, which are considered the most important methodological factors. As to the last element, the toll levels, these are determined once the scheme has been defined in terms of the other variables.

Interday Temporal Toll Structure

According to the basic theory of congestion pricing, travelers must be charged the difference between the marginal and the average cost of their trips, and the social optimum will be achieved. The system is assumed to be in equilibrium before the introduction of the toll policy. As a result of pricing it reaches a new equilibrium, namely the social optimum. If we make the safe assumption that the underlying adjustment process is not instantaneous, we need to be concerned about how to move to the social optimum. There appear to be

---

1Hau (1992) describes numerous pricing schemes in more detail and discusses the entire spectrum of currently available pricing mechanisms such as manual tollgates and electronic toll collection using AVI or smart card technology. The discussion focuses on differences in technology, and quantitative evaluation of the various schemes is based on the cost-per-transaction as a measure of cost effectiveness. A comprehensive discussion of electronic toll collection systems, which are key to efficient implementation of congestion pricing, is provided by Bernstein and Kanaan (1993).
3.1 Characterization of Congestion Pricing and Traveler Information Systems

WHERE

Spatial Structure

Facility-Based
(Orange County)

Zone-Based

Multiple Zone
(Hong Kong, Netherlands)

Single Cordon
(Singapore, Bergen)

Multiple Cordon
(London)

Other
(Cambridge)

Temporal Toll Structure

Interday
Static
Dynamic

Intraday
Single Step
Multi-Step
Continuous

CONGESTION PRICING
SCHEME

Toll Classes
WHO

Toll Levels
Redistribution
Procedure

Other
Enforcement Technology
...

HOW MUCH

Figure 3-1: Congestion Pricing Design Factors
at least two toll policies which may get us there. The first is through charging the “final”
toll, \( T^E \), which is the difference between the marginal cost and average cost at the new
equilibrium. This is the denoted as the \textit{static} approach, where commuters are expected
to respond to the fixed tolls in the same way as they would respond to the other fixed or
constant elements of the network, such as link lengths. For a static structure, the toll at a
particular time of day is constant from day to day of day.

Alternatively, one could (theoretically) charge the true marginal cost for every trip as
it is completed. In this approach, if perfectly implemented, the system would work its way
toward the new equilibrium in the same manner as it was assumed to have reached its
initial equilibrium. Clearly, the first approach is simpler, whereas the second would require
a sophisticated system that can both compute the actual marginal cost of a particular trip
and charge it to travelers as they reach their destinations. We will use the term \textit{dynamic}\(^2\)
to describe all forms of pricing whose temporal toll structure on a given day depends on
the actual flow or congestion levels prevailing at the time.

\textbf{Intraday Temporal Toll Structure}

The intraday temporal toll structure describes the variation of the toll levels with time over
a given day. Much of the literature on congestion pricing that has considered the dynamic
nature of the problem by including departure time decisions has derived the optimal toll as
a static (in the day-to-day sense) quantity that varies \textit{continuously} over time (e.g. Vickrey
[1969], Henderson [1974], Braid [1989]). Singapore and Hong Kong are examples of static
tolls that are \textit{interval-based}, i.e. various forms of step tolls. The simplest step-toll consists
of a single toll period, during which a toll is charged, bordered by “free” off-peak periods
during which charging does not take place, or is carried out at lower rates. This is depicted
in Figure 3-2, where a toll of value \( c_2 \) is charged between times \( t_1 \) and \( t_2 \), and a toll of
\( c_1 \) outside it. The Singapore scheme has a single toll period for the morning commute; at
other times during the morning peak-period, no toll is charged. More complex multiple
step structures can also be devised. The Hong Kong scheme had a three level toll structure,
with peak, intra-peak, and off-peak toll levels.

Dynamic pricing systems are generally expected to result in continuously changing tolls.
The Cambridge scheme is an example, although there are only two toll levels (zero or fixed
toll) for every beacon pair. One can also conceive a dynamic system that is interval-based,
such that the measured levels of congestion translate into costs that are constant over some
duration.

\(^2\)“Dynamic” means different things to different people; even the static toll is dynamic in the sense of being
based on departure times and being time-variant. Perhaps the term “flow-sensitive” is more appropriate,
but it can be just as easily misinterpreted.
Figure 3-2: Interval Based Tolls: A Single-Step Toll
Spatial Structure

For spatial structures, the main choice is that between zone and facility-based schemes. The analogue to the idealized single route toll on general networks is route pricing. In such a scheme, the final equilibrium tolls for each route are calculated, and the commuter is charged each day according to the route traveled. As pointed out above, Dafermos and Sparrow (1969) have shown for the case of route choice only, that it is not generally possible to use link-based tolls that are indifferent to user’s routes or OD’s to replicate the effect of route pricing. Schemes to implement route pricing directly have not been devised yet, although technological advances at least permit their conceptualization. For time-variant tolls the route charges will have to be time-dependent as well. Thus, even if the charging mechanism for implementing route pricing were developed, more work needs to be performed on dynamic network modeling to derive these time-dependent tolls. As a result, approximate ways of implementing congestion pricing have been devised in practice, where only a subset of the possible links/routes are charged.

Outside the US, congestion pricing schemes have invariably been conceived in relation to urban zones, except for the Dutch scheme, where entire regions were to be tolled. In such zone-pricing, vehicles are charged for entering or crossing a zone at certain points. Figure 3.3 shows three basic zone configurations, where O is an origin and D1, D2, D3, and D4 are possible destinations. They are described as follows\(^3\). (i) Single-cordon, where there is a single zone, usually around the CBD, whose perimeter possesses a common temporal toll structure and does not permit avoiding the toll during charging times (Singapore, and Bergen and Oslo in Norway, are examples of such structures.) (ii) Multiple cordon, with several concentric cordons, such that an inner zone cannot be reached without crossing the exterior cordons. These have been proposed for London, UK.; and (iii) Multiple zone, such that no zone’s boundaries are contained in another, an example of which is the Hong Kong experiment. The above schemes are all based on zone boundary crossing at specific times. In contrast, the Cambridge scheme is zone-pricing based on distance and time elapsed between points, and the 1974 proposal for London was based on the presence within the cordon during specified times.

Recently, the emphasis in the U.S. has shifted to facility-based congestion pricing, at least during the initial phase of congestion pricing implementation. The salient feature of this approach is that in contrast to zone-based methods, the option of using an untolled route\(^4\) to the destination is preserved. Changing to another mode is thus not necessary in order to travel during the toll period. Two variants are: Pricing one or more bridge or other

\(^3\)There are no universally accepted terms for these configurations, nor are there agreed ways of classifying pricing arrangements in general.

\(^4\)Or a route not subject to time-variant pricing.
Figure 3-3: Examples of Zone-Based Pricing
Chapter 3  Identifying Candidate Combined Systems

point of access to a CBD (e.g. the San Francisco Bay Bridge), or pricing a limited-access highway or freeway (e.g. the privately proposed Orange County toll roads). The difference is that in the latter approach, there is generally more than one possibly exit point for every entry point. Thus for any vehicle, both entry and exit points must be determined.

3.1.2 Characterization of Traveler Information Systems

Just as there are a variety of different ways to implement congestion pricing, many different types of driver information systems have been proposed. To some extent, these differences are driven by the technologies employed in the particular system. However, for our purposes it is more important to consider the following four functional aspects, shown in Figure 3-4. [For a more complete discussion see Kaysi (1992), Ben-Akiva et al. (1991), and Mahmassani and Jayakrishnan (1991)].

Nature of Information

First, driver information systems can vary in the nature of the information they provide. That is, they can either provide status or guidance information. Status information merely consists of information on congestion levels, travel times etc. without making recommendations as to how to respond to the reported situation. Guidance, on the other hand, consists of specific directives such as which route to take or whether to delay departure for a certain time. In addition, the information they provide can either be (or based on, in the case of guidance) historical, current, or predictive. Historical information reflects trends observed in the past, such as the average travel time on a particular route at a specific time, calculated over some recent period. Current information simply reports on the present state of the system. Predictive information attempts to extrapolate current conditions to the near future. The exact nature of the information will influence how people will respond over both the short run and the long run.

Spatial Structure

Second, the spatial structure of driver information systems can vary in important ways. In particular, they can vary in their geographic extent or coverage. Pre-trip information is received at the trip’s origin, permitting both route and departure time adjustments, while en-route information allows only route adjustments. The specificity governs whether information is provided to individual vehicles (perhaps differing from vehicle to vehicle), or whether the same information is provided to all of the vehicles in a given area (such as occurs with a variable message sign). The spatial structure will largely determine the “resolution” of the ATIS (i.e., how specific the system can be in influencing traffic patterns).
Figure 3.4: Information System Design Factors
Chapter 3  Identifying Candidate Combined Systems

Temporal Structure

Third, different driver information systems can have different temporal structures. Most importantly, the frequency with which the information is updated can vary. This too can have an impact on the resolution of the TIS. In fact, Kaysi (1992) has shown that the temporal structure can have a considerable impact on whether the system will experience overreaction effects.

Provision Criteria

Finally, it is possible to supply information based on a variety of criteria. Of course, the general aim is almost always to minimize some individual’s or group’s travel cost. However, how the group is defined can have a considerable impact on the ultimate operation of the system. Some possible groups include: those vehicles at a particular information provision “point”, all vehicles within 2 kilometers of a particular point, all vehicles currently on the network, and all vehicles currently on the network plus those expected to be on the network in the next \( t \) minutes.

3.2 Desirable Properties of Pricing and Information Schemes

One aspect of the work on congestion pricing and information systems deals with setting criteria that should be met by practical systems in order to achieve the desired ends.

3.2.1 Congestion Pricing

In this section we review several criteria important in the design of a good pricing scheme. General evaluation or design criteria for congestion pricing have been discussed several times in the literature, including the Smeed Report (Ministry of Transport [1954]), Thompson (1990), Hau (1992), Guiliano (1992), and May (1992).

Congestion pricing implementability criteria can be categorized into two broad groups: Prior to implementation, that of acceptability of the scheme, and after implementation, that of operability. Acceptability simply captures the public’s attitudes towards a proposed scheme and is centered on a scheme’s perceived attributes. The persistent rejection of congestion pricing suggests that (i) the perception may differ drastically from the actual attributes, and/or (ii) currently used models and theories predicting the impacts of congestion pricing that have generated enthusiasm on the side of its advocates (such as transportation planners and economists), are inaccurate or fail to address important issues. The work on acceptability consists primarily of identifying possible reasons for public opposition. As such, it is very broad and tends to be rather speculative. Operability includes issues of
3.2 Desirable Properties of Pricing and Information Schemes

effectiveness, safety, and general workability of the scheme. Several specific design criteria are discussed below with reference to acceptability and operability.

Equity

The Smeed Report’s nine primary criteria of a good road pricing scheme included the requirement that the “incidence of the system upon individual road users should be accepted as fair”. Indeed, equity is an important acceptability criterion, and equity problems are often used to explain public opposition to congestion pricing. One should distinguish between the “actual” equity which is the output of policy models, and the perceived equity which represents the subjective impression of the consumer. Equity in congestion pricing is further discussed in Appendix A.

Skepticism

Skepticism with regard to the outcome of congestion pricing schemes derives from the lack of existing schemes which could demonstrate the potential success. The difficulty of understanding the benefits of congestion pricing is another factor - even some insiders have misconceptions about the policy. Moreover, the inherently dynamic nature of congestion pricing makes realistic modeling difficult, and thus the ability of existing models to accurately predict behavioral response is not clear.

Privacy

Privacy concerns formed part of the negative evaluation of the Hong Kong scheme, although advancements in technology have made this much less of an issue. Today, electronic toll collection systems of various specifications (including Read/Write and Read Only systems) are widely used in the US and other countries (see Bernstein and Kanaan [1993]), primarily on major highways, and the privacy issue has not emerged as a major obstacle in those cases. Of course the Hong Kong system, which was based on logging travel at numerous points within an urban area, rather than freeways, provided a more detailed record of a person’s travel. Yet a particular system can be adapted to address public concerns regarding privacy by devising rules regarding the use and maintenance of the collected data. In this regard the situation appears little different from the itemization of telephone calls,

Redistribution

Hau (1992) argues that a successful scheme must provide for redistribution of revenues in order to stand a chance of acceptability in a democratic society. The basic economic theory of congestion pricing assumes that the revenues that are collected constitute a pure transfer
payment and are thus not considered a loss to society. In practice, the problem is more complex. Some of the revenue is offset by operating costs. More important, there is a broad array of proposals on how revenues should be channeled back to the public.

Goodwin (1990) suggests a simple “one-third” rule, where equal amounts are spent on highway improvements, mass transit, and general tax relief or increased general expenditures. Small (1992b), in the most detailed analysis of redistribution, criticizes the share to be spent on highway improvements as too high on grounds that (i) the scope of government would be expanded in a manner that suggests taking sides in a divisive debate, namely that over supply vs demand based relief measures, and (ii) the presence of congestion pricing would at least partially substitute for highway-based capacity expansion. He also reviews the points of view that (i) the only politically salient case for revenues is to fund new highways (Gomez-Ibanez [1992]); and that revenues from each corridor be targeted to highway improvements in that corridor (Bay Area Economic Forum [1990]).

Criticizing that limited approach, he instead proposes, for *area-wide implementations* his own one-third rule which divides into: (1) monetary reimbursements to travelers as a group, in the form of funding employee commuting allowances, and the reduction of road user taxes, (2) substitution for general taxes now used to pay for transportation services through reducing sales-tax surcharges in the affected areas and by rebating a portion of property taxes, and (3) new transportation services, through highway capacity, mass transit improvements, and improvements in transportation-related facilities and services in business centers. This arrangement is based on a principle advocated by Burntaw (1991), which says that losers from environmental policy decisions should be compensated by means of “linked compensation”. This means that losses are offset by measures that directly alleviate the harm done, which is claimed to be fairer and more understandable than monetary transfers by most people. Small also presents a case study of Southern California, based on the assumption that all congested freeways and arterials in the five-county Los Angeles area would be subject to congestion pricing. Using strong simplifying assumptions, actual revenues and their allocation to the various components of a redistribution scheme are presented.

"Accurate" Charging

The Smeed Report states that charging should be closely related to the amount of use made of the roads. Similarly, Hau (1992) suggests that to enhance economic efficiency, the system should be able to charge directly, as closely as possible, the external costs arising out of road use. These recommendations for *directly use-based charging* may have their root in the formulation of the classical congestion pricing model, which says that users should be charged their true total cost. The practical implication of this criterion seems to be a form
of dynamic charging. This is an operability issue, since efficiency is involved, as well as the workability of the scheme. We will discuss the potential acceptability problems of this criterion in the next section.

Simplicity and Transparency

The Smeed Report postulates that the method should be simple for road users to understand. Hence, extremely complex pricing gradations should be avoided (Hau [1992]). The system should also exhibit transparency, such that the user should be able to determine a trip's potential cost before making the relevant decisions. In other words the toll-route matrix should be available to the traveler. These issues relate both to the efficiency (operability) of the scheme, and to its acceptability.

3.2.2 Information Systems

Evaluation criteria for traveler information systems have received less systematic attention than for congestion pricing.

Validity or Reliability

Ben-Akiva et al. (1991) and Kaysi (1992) point out that if the supplied information provided by a route guidance is of low validity (i.e. differs significantly from its actual values), it is likely that users will in due time abandon the system and follow their own intuition. This is a post-implementation issue. Related to it is the efficiency aspect of a system with low validity, since one would expect efficiency to increase with validity. Kaysi (1992) points out that none of the existing projects, nor any of the theoretical research, has considered the response of motorists to route guidance in setting such guidance. The main conclusion of that research is that, for guidance systems, the provided information should be based on predicted future traffic conditions that takes the user response to information into account, since otherwise serious errors may occur. For real-world networks the current modeling and computational capabilities are judged not yet advanced enough to make the necessary accurate prediction.

Equity

A (pre-implementation) acceptability constraint is associated with the incidence of benefits and costs on the various user groups: If the information system is of the type that will be owned by only a subset of the commuting population, equity concerns may be a major obstacle to implementation of a system, since it is likely that, in general, equipped users will benefit and unequipped vehicles lose as a result of implementation (Mahmassani and
Chapter 3  Identifying Candidate Combined Systems

Chen [1991]). It is true that this is conceptually similar to the acceptability problem of congestion pricing: Not only can the affluent get a better product, they will do so at the expense of the less well-off. The ease with which pilot projects have been implemented around the world, and the fact that no behavioral adjustments are required by the lower income groups, indicate that the perceived inequity in this case is likely to be lower.

Spatial Structure

The choice of geographic extent and specificity depends to a large degree on the fundamental objectives of the system. The specificity of a system that must be purchased individually will be determined by the market, and depends, among others, on price and quality. One relevant indicator of quality for market penetration is the magnitude of the benefits that accrue to equipped vehicles. It is interesting to note that the modeling inadequacies of currently designable systems generally show that total benefits peak at low penetration rates of equipment. “Public” systems are likely to extend over larger areas, and are potentially accessible to all within the area, such as radio broadcasts or variable message signs.

Temporal Structure

The temporal toll structure should be as high as possible for maximum responsiveness and accuracy. The choice of which group’s costs to minimize can be illustrated best by taking the two extremes, namely an individual driver, or the entire relevant traveling population. With a single driver, the computational effort should be comparatively low, and, with sufficiently frequent temporal updates, the perceived equity/effectiveness problem that may arise when two vehicles with similar destinations that are close in space and time yet experience widely differing travel times as a result of guidance is likely to be minimized. On the other hand, providing guidance with the goal of minimizing total travel cost is likely to lead to unavoidable disparities in the travel times of such similar vehicles, and if observed by the commuters, may cause acceptability or abandonment problems.

Some of the implementability issues discussed above are straightforward in terms of their implications, such as the temporal update frequency. Others, such as specificity, exhibit trade-offs that have not been explored fully, and depend on the basic objectives of providing information systems.

3.3 Selecting A Congestion Pricing Scheme

This section seeks to identify, from among the many possible congestion pricing schemes, those that stand a good chance of successful implementation. The discussion will include elements of the criteria discussed above, and address the main methodological design factors,
3.3 Selecting A Congestion Pricing Scheme

namely the temporal and spatial toll structure. The various possibilities referred to here are those that were described in Chapter 2. We also decide on the basic structure of the information system to be analyzed. The arguments made below reflect our personal viewpoints, and due to the absence of more practical experience with different pricing schemes, can certainly be debated.

3.3.1 Temporal Structure

Two main decisions must be taken with regard to the temporal toll structure: The within-day variation, and the manner in which the tolls are calculated, i.e. the day-to-day variation. We must decide whether to opt for continuous or interval-based charging, and whether charging should be static or dynamic. After looking at these two elements separately, we evaluate the combination considered most appropriate, namely static step-tolls.

Interval-Based vs Continuously Varying Tolls

Optimal congestion tolls are continuously time-varying quantities. Real-world application of such a toll has been recognized as impractical, usually because of the difficulty of administering such a toll. The problem with optimal tolls, even if they could be administered in a time-varying fashion, is that the resulting “price schedule” appears far too complex to result in the travel behavior changes that the pricing scheme is supposed to elicit. van Vuren and Smart\(^5\) (1990) note that continuous variation of tolls overburdens drivers with the amount of information they require for optimal decision making and thus recognize that a limit exists to the extent to which prices may be varied. Moreover, natural variability in travel times make tolls that vary precisely by the minute impractical. van Vuren and Smart (1990) also argue that, for the adjustment of users to the new situation to proceed smoothly, the levels of tolls have to be stable over time, and thus need to represent averages corresponding to the “conditions most likely to prevail in a particular area at a particular time”.

In practice, therefore, a form of step-toll or interval-based toll has most commonly been used or proposed. This is the case in both Singapore and Hong Kong discussed above, and is also the case for public utilities. In the telephone industry in the U.S., time of day pricing is usually based on three pricing periods, with constant rates for each period. Park and Mitchell (1987) studied optimal time-of-day measured rate prices for local telephone calls. They denote a feasible pricing structure as one that is “simple enough for the consumer to grasp”, arguing that feasible tariffs are perhaps limited to three price periods that repeat

\(^5\)A similar argument was made by Altshuler et al. (1979), and Vickrey (1980) for the case of transit pricing.
from day to day. The Washington, DC, subway fare system consists of a single peak-period in the morning and another in the evening. Utility companies have up to four similar predetermined periods, with fixed price rates within those periods.

**Dynamic vs Static Tolls**

In our view one major disadvantage of dynamic pricing that is sensitive to the actual levels of congestion, in contrast to a static structure which does not vary from day to day, is that users cannot know in advance whether and what they will be charged for a trip because of the stochasticity inherent in both recurrent and, in particular, incident congestion. This is very unlikely to be politically acceptable, since one is charged an unpredictable amount, which can be very frustrating. For example, an unpleasant situation could arise where a commuter could find himself, through no fault of his own, in a traffic jam caused by, say, an accident, and, in addition to the excessive travel times, face a monetary penalty! Oldridge (1993), defending the design of the proposed Cambridge system against criticisms to that effect made by van Vuren (1992), argues that this is less of a problem since travel times at various times of the day are relatively constant. Of course, in that case there is no need to make the charge between beacons sensitive to the actual travel time.

It is also likely that dangerous driving will be used to "beat" the system, since a system that charges for the time spent in traveling between two points clearly encourages faster driving. Moreover it may not be possible to avoid the toll by early departure if there are so many early departures to result in shifted congestion. Therefore, the predictable effect on travel patterns (e.g. route and departure time choices) is lost. Although it is not possible to rule out improved efficiency for an implemented system (assuming that it has overcome the acceptability obstacle), it is certainly difficult to predict the efficiency impacts. Therefore, real-time tolls that change on a daily basis, whether interval-based or continuous, in ways that are unpredictable and unknown to the commuter in advance suffer both from serious acceptability problems, potentially poor safety, and highly uncertain outcomes.

Static tolls whose value at a certain time of day and a specific location remains constant over time do not suffer from the above problems. Travelers will know what they must pay at various points of their trip. Combined with an interval-based within-day structure, trip planning should be simple, and more confidence can be placed into models that predict user response. However, it should be noted that an attribute of interval-based static toll structures is that they are not automatically responsive to changes in the underlying demand patterns - since peak-periods are really only meaningfully defined in terms of the desired arrival times which commuters target in our context. In particular, it is to be expected that employers will coordinate with their employees more flexible schedules - thus the pricing scheme could in effect serve as a device for implementing staggered work hours and/or flex-
time. This may well be one of the declared objectives of the pricing scheme, and would reduce the scheduling costs of those who as a result of the toll incurred greater such costs. This is in contrast to the theoretical method of charging directly the marginal cost of each trip, which does not suffer from this drawback. Therefore, the interval-based toll should really be quasi-static, so that network conditions and demand levels can be periodically re-evaluated and the tolls appropriately adjusted. We will refer to interval-based, quasi-static toll structures as step-tolls for simplicity.

**Drawbacks of static step-tolls**

Although we have chosen step-tolls as the most suitable temporal toll structure, we realize that they are not perfect. Some of their potential implementability problems include safety, crowding, and uncertainty, and are discussed below.

**Safety** Although (static) step-tolls have several advantages over dynamic tolls, the discontinuities in travel cost as a function of time give rise to various potential safety and efficiency problems. One such problem is the possible safety hazard that arises when commuters attempt to get through the toll station before the toll period begins. This problem can be completely mitigated only by locating the toll stations at a commuter's home - a fictitious concept, but one which would incidentally solve the problem associated with travel time variability discussed in the next subsection.

**Crowding** Another problem is caused by crowding effects likely to be seen when implementing step tolls. It is possible that some commuters will leave their homes in such a manner that they will arrive at the toll station just after the end of the toll period - furthermore, few commuters would be expected to pass the toll station just before the end of the toll period. Thus, to arrive after the toll period would mean not only not paying the toll but experiencing reduced congestion as well. In practice, commuters will respond to this potentially "very good deal" by crowding at the time when the toll period ends - both in terms of the timing of their departures, and by way of waiting, if possible, in the vicinity of the station if they should arrive there sufficiently close to the end of the period. From an acceptability point of view, this type of system may be problematic, since the crowding may be anticipated and perceived as a possibly very negative experience (e.g. similar to shopping on Christmas eve) by those who would be intent on traveling after the toll period ends, and as inefficient by others. The severity of this problem will depend on the specifics of a particular implementation.

**Impact of Travel Time Variability** A further important concern related to interval-based tolls arises when considering the variability in travel times that occurs in transporta-
tion networks. Let us illustrate this by way of an example. Suppose a commuter lives 20 miles from the point at which he is charged his single congestion toll. He uses a particular route to get to work, and although he may occasionally change his departure time from home and/or his usual route, he decides today to follow his routine travel pattern. The toll period extends from 7:00 a.m. to 9:00 a.m., and he regularly passes the toll station at 6:55 a.m., paying no toll. However, due to an incident, or perhaps simply due to unexpectedly heavy traffic, the trip from home to the toll station takes five minutes more than usual - and our commuter, who hates to pay the toll, has to do so today, resulting in one frustrated citizen.

In practice, he may be expected to get into such a situation only a few times after the implementation of the congestion pricing scheme. It is likely that he will alter his travel pattern, perhaps by leaving sufficiently earlier, to make sure he passes the toll collection point before the toll period even in the case of heavy traffic. This, apart from its potential as a continuous source of frustration, may also be inefficient. The experience of the Singapore congestion scheme might provide useful empirical evidence on the user response to this feature of pricing schemes.

One method to overcome this problem would be the installation of local toll card readers sufficiently close to commuters' homes so that punctual arrival at those readers would be easily achieved. The toll would then be charged when passing the "real" toll station further down the route based on the time the user passed the first card reader. This would require a somewhat more involved system, but does not appear infeasible even with today's technologies.

3.3.2 Spatial Structure: Zone vs Facility Pricing

The selection of the spatial structure of the scheme is a high-level choice which will be based on the particular perspective of what the role of congestion pricing should be and how people respond to pricing.

The most popular spatial arrangement seeks to charge, in one of a number of ways, travel within a zone, usually the core of a city. As we have pointed out, the main options are single-cordon, multiple-cordon, or multiple-zone schemes. Optimal zone boundary location (even for a single zone scheme) is a currently intractable problem if pursued in a time-variant setting. In practice, city structure may suggest layouts that are intuitively optimal or near-optimal, or that would reduce implementation costs. For example, downtown areas that are accessible only by a handful of bridges may make those bridges the convenient and economical choice.

The other option is to implement congestion pricing on select (limited-access) facilities. This approach is seen as a (partial) answer to the question of how to go about introducing
3.3 Selecting A Congestion Pricing Scheme

one or more pioneer schemes that, if successful, will serve as landmarks that will help to popularize the concept. Suggested facility-based pricing strategies to overcome the acceptability problem that are widely believed to stand the greatest chances of success include (Poole [1992], Small [1992a]):

- Introducing congestion pricing on new highways or bridges that are already planned as toll facilities

- Adding time-of-day variation to the toll structure of an existing toll bridge or other bottleneck, possibly coupled with introducing electronic toll collection technology.

- Opening existing or planned high-occupancy vehicle lanes to non-qualifying vehicles willing to pay a fee.

Pricing limited-access highway facilities in particular is very flexible in that there is a choice of where to charge users. Two obvious candidate points are those of entry and exit. Since currently employed tolls on such facilities are time-invariant and distance-based, commuters are charged at the exit based on their entry point. With time-varying congestion pricing one must decide whether the toll charged will be based on the time of entry to the facility or the time of exit (or the time of passing an intermediate point). If pricing is based on the time of exit, acceptability problems may occur as a result of the increased unpredictability of travel costs when travel time variability is high between entry and exit.

There are two ways of looking at facility-based pricing. According to the first view, it is only the first step in moving towards area-wide schemes. Thus, if the introduction of congestion pricing on the chosen facility goes well, more facilities can be priced, until the desired effect is obtained. This makes most sense if the basic objectives for the adoption of the policy are related to the mitigation of one or more problems caused by excessive congestion, such as high air pollution levels. The second view of facility-based pricing is more limited in scope, namely to modify the demand pattern on the facility with little regard to the remainder of the network. An obvious reservation in that case is the potential for spill-over effects that would send traffic to other parts of the network that are not priced (or not priced with peak/off-peak differentials).

3.3.3 Selecting a Scheme

With regard to the temporal structure of a toll scheme, we have argued that (quasi) static pricing is to be preferred to dynamic pricing for reasons of acceptability and efficiency. Similarly, an interval-based toll was preferred over a continuous one for reasons of simplicity, acceptability, and predictability of impacts.

Yet even for a quasi-static, interval based temporal toll structure, the spatial toll structure must be carefully examined to determine whether its combined effect with the former
results in a still manageable time-route toll matrix. That is, a commuter has a set of possible routes to work, and for every route, the effect of temporal and spatial toll structure results in both a total cost and a cost at certain points along the route that is time-dependent. For a (hypothetical) transportation network with zero variability in travel times, these two variables are unique for every departure time, but even so the functions expressing these relationships can become rather complex for multiple zone or cordon schemes, and the departure time intervals over which these variables are constant can become quite short. Furthermore, for a given origin and departure time, different routes may have different time-dependent toll structures, potentially causing bewilderment to the traveler.

For example, the Hong Kong scheme was designed as a multiple zone system in such a way that some travelers would have to traverse several zones during their trips. Some commuters were faced with a potential of nine different trip charges, depending on the arrival time at the various toll stations (zone boundaries). In our opinion this exceeds what can be considered reasonably simple. Even if such a scheme is found acceptable, and satisfies operability criteria, (people may get used to it; note that this aspect of the Hong Kong scheme has not been raised as problematic in the discussions following its defeat as a general policy), it appears unreasonable to develop a toll scheme with a "sophisticated" toll structure such as this since the assumptions of rationality inherent in the modeling process become unreasonable when the commuter is faced with a situation of great complexity and increasing unpredictability.

We thus argue most strongly for interval-based, static tolls. With regard to spatial structure, there is no obvious preference between the different zonal arrangements as long as the overall design takes into account the interaction between the spatial and temporal structures described in the preceding paragraphs. Facility-based pricing is considered as a phase-in approach.

For the purpose of the analysis presented in this research, facility-based pricing will be emphasized rather than zone pricing because of its imminent implementation in the U.S. Although the modeling of the two approaches does not differ significantly at the level of detail at which it will be pursued in this research, special features relevant to facility-based pricing will be given emphasis. The temporal toll structure is quasi-static and interval-based. These include most prominently the issue of toll station location, which will be explained further and analyzed quantitatively in Chapter 5. The next two sections similarly describe some evaluation criteria for information systems and combined systems, and concludes with a discussion of our chosen combination vs others that have been suggested elsewhere.
3.4 Selecting an Information System

Our aim in this research is to explore the combined implementation of congestion pricing and traveler information systems. Little disagreement exists among information system developers about what constitutes a good information system. The most important criterion for a good information system is reliability, i.e. the capacity to provide information or guidance which is accurate and thus makes users confident that believing or following the information system is to their advantage. Differences exist more on how to model various aspects of the interaction between the system and the user, or on the nature of the prediction algorithms used.

Both pre-trip and en-route information systems are being extensively researched. However, en-route information has received more attention at this point in the form of the deployment of route guidance systems. Modeling en-route is "simpler" compared to pre-trip information which can affect not only route choice but also departure time and mode (including the possibility of deciding not to proceed with a trip). For simplicity we will model an en-route guidance system. With regard to the nature of information provided, the advantages of guidance over status systems is that the former takes the work out of making a decision. However, commuters may not wish to be told what to do. More important the guidance is likely to be less accurate the more complex the situation. In that case commuters may make less accurate decisions than the control center, but a poor decision cannot be blamed on the guidance provider. We select a vehicle-specific guidance system so as to reflect the major effort directed at the design of in-vehicle information systems.

The type of information provided will be based on minimum cost for the commuter(s) currently receiving information system located at a particular node (or nodes). Theoretically such systems may produce less benefits than a system which provides information so as to minimize the average costs for a larger group of people. However providing such "system-optimal" information is likely to be even more difficult to achieve. Furthermore it can be expected to involve providing information to some individual travelers that will be sub-optimal from their perspective, and may thus quickly become unpopular.

We have now chosen both the congestion pricing scheme and the information system, based on arguments that considered those policies independently of each other. To decide whether those systems, when implemented together, still result in a good system, we return to the literature of joint implementations reviewed in Chapter 2, and examine it in light of the evaluation criteria developed above.
3.5 Selecting the Joint Pricing/Information Scheme

In the previous two sections various design aspects of pricing and information systems have been discussed. We suggested a particular form of pricing, with an emphasis on it being the most suitable from an implementability point of view (quasi-static, interval-based temporal toll structure) and being representative of current efforts in the US (facility-based pricing). This section discusses whether integration of the two policies necessitates a fundamental change in one or both policies by means of a discussion of some proposed combined systems, which have been referred to in the previous chapter.

3.5.1 The Role of Congestion Pricing and Information Systems in Congestion Alleviation

When designing an integrated system it is important to distinguish between two types of congestion, non-recurring and recurring. Non-recurring congestion, which does not follow a regular day-to-day pattern, is due largely to incidents (broadly defined) occurring in the system. Recurring congestion follows a regular pattern, and is due simply to the "normal" dynamics of traffic flow. It is generally believed that an integrated system should reduce both types of congestion if possible. There are many different ways that an integrated congestion pricing and driver information system can approach this problem. Congestion pricing can be used to manage recurring congestion and/or non-recurring congestion, and the information system can be used to manage recurring congestion and/or non-recurring congestion.

Above we have argued, through our choice of systems, that it is best to use congestion pricing to manage recurring congestion, where it has demonstrated clear potential for success. Its application to reducing incident congestion would require some form of dynamic tolls, which are marred by problems of acceptability and uncertain efficiency gains. Information systems, on the other hand may help to reduce recurrent congestion by educating travelers about traffic on unused routes, which may bring about a move to a point at which total system costs are lower. Once commuters are familiar with the available alternatives, it appears that information systems will be most useful in managing non-recurring congestion.

It is also assumed that non-recurring (incident) congestion will be reduced by information systems, which are assumed not to affect recurring congestion directly. Thus adjustment to sudden changes in network capacity is based on the supply of information about alternative routes. It is conceivable that the default or average behavior is based on the network behavior in such a way that incident congestion is taken into account, for example through a dependence of equilibrium flow patterns on the expected network capacities, rather than incident-free default capacities. An incident management system that reduces the sever-
ity and duration of the network capacity reduction might thus have a direct impact on the default behavior that underlies non-recurring congestion. In this study, non-recurring congestion is modeled in the form of temporary reductions in network capacity caused by incidents. The details of the equilibrium model, and the method of modeling incidents, are described in Chapter 4.

3.5.2 Drawbacks of Other Combined Systems

Most of the discussions of joint implementations of pricing and information systems have visualized situations in which prices would be used to engender compliance with the information provided. Thus, de Palma and Lindsey (1992) employ the concept of state-dependent tolls, which vary dynamically as capacity changes in the network (recall that capacity is assumed to be constant on a given day but is allowed to vary stochastically from one day to the next). Smart and van Vuren (1990) conceive a combined system that uses pricing as a signal for appropriate route choice. The Cambridge system is contemplating coupling an information system to their proposed dynamic tolling scheme. In the telephone industry, Park and Mitchell (1987) conceive a system where a user would call to receive a quote on the current price of placing a call.

There are several potential problems with such schemes:

- In addition to the frustration of dealing with a system that is unpredictable in terms of costs, as we have pointed out above, it is not clear which prices would result in the desired compliance.

- The current state of the art indicates that accurate information systems are not yet ready to be implemented. It seems only to compound the problem if we attach financial consequences to not complying with information directives.

For these reasons we do not believe that the integration of pricing and information should proceed by making prices dependent on real-time congestion, or that information should be coupled to tolls, as has been suggested by some of the cited research. However, the integration of the two systems spawns certain implementability issues which, while not altering the fundamental role the systems should play with regard to the different types of congestion, must nevertheless be addressed when designing a successful combined system. In the next chapter we outline some of these issues, followed by a detailed description of the model used to examine them.

---

6 This does not seem to be an improvement from the perspective of user-friendliness, either. It is not at all clear when to expect to find the cost to be as low as desired, nor, if a call revealed a rate that is too high, when the cost will drop to the desired value.
Chapter 4

Modeling Combined Pricing and Information Systems

We assume in this research that pricing affects average or default behavior, which is described by an equilibrium travel pattern. Incidents in the system, such as traffic accidents, reduce network capacity, and information on the congestion caused by these incidents may be used to adjust route choices. The main objectives of this chapter are to discuss some of the design issues that arise from the integration of pricing and information schemes, and to describe the model used to investigate some of these issues in a quantitative setting.

After reviewing integrated design issues, we justify the decision to analyze simple rather than general networks. This is followed by a description of the network and the equilibrium model used to capture default behavior. Subsequently we describe the modeling of incidents, and the information system, the assumptions about traveler response, and the simulation model used to implement the combined pricing and information system.

4.1 Integrated Design Issues

Some of the more important integration issues specific to the congestion pricing and driver information systems under consideration here are summarized in Table 4.1. In this table we distinguish between temporal (i.e., how the systems are designed to behave over time), spatial (i.e., how the systems are designed to behave over space), and demographic (i.e., how the systems are design to behave across different people/vehicles) issues.

One of the most important integrated design issues is that route guidance can no longer be based on simply the travel time (or distance), but must in some way take the route and time dependent toll costs into account, an issue alluded to by Brett and Estlea (1989). If only traffic information is provided, it is possible to continue to provide the information on travel times or distances and simply add the relevant route-toll combinations; this could potentially
Table 4.1: Specific Integrated Design Issues

<table>
<thead>
<tr>
<th></th>
<th>Info. Systems</th>
<th>Congestion Pricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal</td>
<td>Info. for bulk</td>
<td>Trapezoidal toll</td>
</tr>
<tr>
<td></td>
<td>Info. during gap</td>
<td>Common end times</td>
</tr>
<tr>
<td>Spatial</td>
<td>Location of info.</td>
<td>Location of toll stations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Common tolls</td>
</tr>
<tr>
<td>Demographic</td>
<td>Penetration rate</td>
<td></td>
</tr>
</tbody>
</table>

become rather complicated. Similarly, route guidance could either provide guidance on the basis of time or distance, with toll information supplied separately, or provide guidance based on total generalized cost. Although the latter is simpler, it must be based on a weighting of time and tolls, which necessitates knowing the traveler-specific values of travel time. This may be a very difficult problem to solve accurately.

For the purpose of this research we will combine the travel time and toll costs (together with schedule delay) and provide travelers with a single generalized cost. In practice this requires that we know commuter-specific parameters to be able to make these tradeoffs. In addition even if a person would in general combine these various costs in a certain manner, it is possible that at the time of receiving information on routes these preferences change. In other words it appears far from clear that these parameters are constant for a particular traveler.

The second specific design issue to consider is whether information should be provided to the bulk. Quasi-static interval-based tolls will likely give rise to various edge effects. In particular, the end of the toll period(s) will probably see some congestion, perhaps in the form of a departure of a large number of people (“bulk”) over a very short time period, that will compensate for the large difference in cost that would otherwise exist between passing the toll station just before and just after the toll. The provision of information to a large number of travelers over a very short period of time may mar the workability of the system, since it may require very frequent updates of the information in order to avoid overreaction. In addition, there could be potential safety implications, if guidance to the bulk results in a significant increase in maneuvering\(^1\). We will investigate this issue in the next chapter.

A third and closely related design issue is whether a true interval-based toll or a “trapezoidal” toll should be used. Observe that the potential edge effects just described are a result of the discontinuous change in cost (over time). A “trapezoidal” toll could mitigate these effects by gradually raising and lowering the toll during a short period before and

\(^1\)Many related questions arise with respect to the gap in traffic that is likely just after the toll period begins.
after the toll period. This could imply that (i) there would be less incentive for dangerous driving and (ii) the bulk would distribute itself and result in less congested conditions. But such a toll structure is potentially very complex, and thus may be viewed with skepticism; in addition, it may not be efficient to provide guidance during those periods. We will not consider trapezoidal tolls further in this thesis.

A fourth design issue is whether different roads should have common toll period end times. With two or more tolled routes, people may think the system is unfair if the toll periods end at different times. Unequal end times may also result in a great deal of route switching when information is provided (because, after an incident, the route with the earlier end time will be much less costly than the route with the later end time), with the obvious resulting safety and workability implications. However, common end times could reduce the efficiency gains from pricing.

Similarly, it is important to consider whether multiple routes should have the same toll. This is important particularly since it raises the question of whether guidance should be based on cost, travel time, or both. With equal toll levels this is clearly less of a problem, because the tolls effectively cancel each other out. But with unequal toll levels, and in the context of facility-based pricing (which retains untolled alternatives), this question must be addressed. This and the previous issue of equal toll periods are examined in the next chapter.

A sixth design issue is that of toll station location. Facility-based congestion pricing will probably be applied to limited-access highways. In the presence of time-varying congestion pricing one must decide whether the toll charged will be based on the time of entry to the facility or the time of exit. The effective toll station location is especially important in view of the fact that incidents may occur and vehicles may be advised to switch routes. With information, those who may be "drawn into" the toll period as a result of an incident but who would otherwise avoid it, can make appropriate adjustments. Will this seem unfair to those with no access to information? Moreover, those with no access to information may adopt very conservative travel patterns to avoid such situations, leading to inefficiencies. We will investigate some of the impacts of toll station location for step-toll congestion pricing in general in the next chapter.

Another design issue is related to the location of the information system. The most important question here is whether, in order to make the integrated system appear to be one "product", it is necessary that the information be provided close to where the tolls are being charged and vice versa. (This issue may also arise with respect to the temporal coordination of the two systems). It is also important to ask whether information that is provided at a considerable distance (upstream) from the toll station will be believed. Further, it may be useful to provide status information (particularly about tolls) even at locations where no en-route adjustments are possible simply to ease drivers’ minds about
Chapter 4  Modeling Combined Pricing and Information Systems

the toll they are going to pay.

The eighth issue is that it is important to observe that there are many ways to affect the penetration rate of information systems (e.g., price). Hence, this raises the question of whether there should be a specific penetration rate goal. With a low penetration rate it may be difficult to sell the integrated "product" to the public (since it will essentially be viewed as a congestion pricing program). In addition, the penetration rate could have an impact on many of the operability concerns raised above. The last two questions are not addressed further in this thesis.

Finally, the improvements (according to some measure of effectiveness) of a combined system could differ from the sum of the improvements obtained by implementing each policy alone. We will test for "superadditivity" in the next chapter.

Though this is not a comprehensive list, it does illustrate many of the issues that must be considered when designing an integrated congestion pricing and driver information system. As stated, some of these issues are quantitatively explored in the next chapter. The remainder of this chapter describes the modeling framework in which the quantitative analysis is carried out.

4.2 Simple vs. General Networks

The quantitative part of this research is limited to the analysis of a single simple network. Before proceeding, we would like to defend this choice vis-a-vis alternative ways to analyze the issues at hand, in particular, the use of general networks.

- The problem of analyzing the joint interaction of pricing, incidents, and information, is very complex. The previous literature on the topic has dealt with highly abstract analytical models that fail to capture essential features of the problem. In spite of certain obvious limitations of the simple model used in this research, we feel that we have moved the analysis to the point where useful conclusions can be drawn.

- Many interesting issues can be studied using the simple model, as we intend to demonstrate in the next chapter, and we feel it is appropriate to understand the interactions in such a setting before moving on to more general and complex, (and as a result, more approximate) analyses.

- An important point to emerge from the recent (and continuing) debate on how to implement congestion pricing in the US is the (nearly) unanimous view that facility-based pricing stands the highest chance of gaining public acceptance. When the intended facilities are limited-access highways, a simple model is less restrictive when applied
4.3 Network Structure

Figure 4-1 depicts the network under consideration. Commuters travel from A to C and have the choice of two routes, using either links 3–1, or links 3–2. The free-flow travel times on links 3, 1, and 2, are 15, 12, and 18 minutes respectively. There are potential bottlenecks on links 1 and 2, just before C. These are modeled as deterministic queues, with capacities of 780 and 300 vehicles/hour respectively. Toll stations are located either at the entrances of links 1 and 2 (at B), or just after the bottlenecks on those links (at C). The OD-flow from A to C is 2200 vehicles. The capacity of link 3 is set arbitrarily high, and hence we have in effect a two parallel route network, with route 1 consisting of links 3 and 1, and route 2 consisting of links 3 and 2. Modeling links 1 and 2 as a combination of a free-flow component of fixed cost and a bottleneck of variable cost (depending on queue length) is convenient and has some empirical validity (Lisco 1983), as such behavior has been observed on urban freeways. In addition the fixed cost component brings the travel time function on a link closer in shape to that of the popular speed-density curves (see also Bernstein and Smith [1994]). Thus $T_1^f = 27$ min. and $T_2^f = 33$ min. The desired arrival time, $t^*$, is 8.0 hrs.

Below we describe the equilibrium model used to represent default travel behavior, followed by an account of how incidents are modeled, how the information system is assumed to work and how users respond to it, and conclude with a description of the simulation model used to examine the pricing and information systems together.

4.4 Modeling Default Behavior using an Equilibrium Model

We assume in this research that there is an average or default travel behavior that is captured by an equilibrium flow pattern. In other words, short-run average flow patterns are assumed to correspond to such an equilibrium. As a result of imposing a congestion pricing policy, users behavior will adjust until a new equilibrium is reached.

4.4.1 Review

The Wardropian equilibrium for static networks is defined as the route pattern such that (1) the cost on all used routes is the same, and (2) the cost on all unused routes is greater than or equal to that on the used routes, for a given OD-pair (Beckmann et al. [1956]). A similar definition is that of user equilibrium, which states that, at equilibrium, no traveler
can unilaterally improve his or her travel time by switching routes (Dafermos and Sparrow [1969]).

These definitions have been extended for the dynamic case, where departure times can be adjusted as well as routes. Thus, a dynamic route / departure time user equilibrium can be defined as a travel pattern in which no traveler can improve his total travel cost by unilaterally switching to a new departure time and / or route (see Bernstein et al. [1993], Jauffred [1993], and Friesz et al. [1993]). Dynamic equilibrium models for general networks are still under development. As an alternative to equilibrium models, dynamic models have been proposed that are based on the notion of a stochastic process in which the system moves from one state to another in a random fashion (e.g. Cascetta [1991]).

The question of whether equilibrium models accurately represent travel behavior has been addressed empirically by a number of researchers. For the case of static equilibrium, validation studies by Florian and Nguyen (1976), Bovy and Jansen (1981), Florian and Tremblay (1986), and Janson et al. (1986), have shown reasonable agreement between traffic counts and flows predicted using user equilibrium models.

We think that assuming an equilibrium pattern as the underlying default behavior is a reasonable approach that has some empirical support, is widely used, and for which no clearly superior (both in terms of reflecting actual travel behavior and computational feasibility) alternatives exist. Since our objective is to examine the effect of congestion
pricing on the temporal demand pattern, i.e. since it is an inherently dynamic policy, we employ a model that is based on dynamic equilibrium.

4.4.2 Model Description

We shall use the deterministic analytical approach developed by Arnott et al. (1990a, 1990b), which is in turn based on a model by Vickrey (1969), to derive equilibrium patterns for networks with and without tolls. We chose this model because it analyzes the optimal step-toll, which we prefer over the continuously varying "fine" toll (see Chapter 3).

Model Basics

As noted in Chapter 2, an essential feature of these models is the notion of a trade-off between travel time and schedule delay. Those commuters who arrive at their destination at, or close to, the desired arrival time, denoted by $t^*$, will have longer travel times than those who arrive early or late. In the particular model adopted, the travel cost function is linear and may be written as

\[ C(t) = \alpha \cdot \text{(travel time)} + \beta \cdot \text{(time early)} + \gamma \cdot \text{(time late)} + \text{Toll} \]  

where $C(t)$ is the cost for a commuter leaving at time $t$, and $\alpha$, $\beta$, and $\gamma$ are the shadow values of travel time, time early, and time late respectively. These parameters have been estimated by Small (1982) with $\alpha = 6.40$ $$/hour$, $\beta = 3.90$$$/hour, and $\gamma = 15.21$$$/hour, and are the values used in this work, assuming a homogeneous population of travelers.

The model assumes that demand for travel is fixed, and commuters can adjust only routes and departure times in response to changing policies and conditions. The travel time on a route $i$ is the sum of the total time spent traveling at constant velocity along the route, $T_i^f$, plus the total time spent queuing at the bottleneck of the links comprising the route, $T_i^q$. Note that in the case of the network used here, only one link on a route is subject to bottleneck congestion. The bottleneck at the end of a link is represented as a deterministic queue. Thus, if the arrival rate of the bottleneck exceeds the capacity or service rate $s$ of that link, queuing will occur. We will outline the main results for the no-toll equilibrium and the step-toll equilibrium below. In both cases we first discuss the results for a single route, followed by the results for the case of two parallel routes. In the latter case we will use subscripts $i = 1, 2$, which are left out when discussing the single route results.

\textsuperscript{2}We detected an error in Arnott et al. (1990b) and provide the correct results here; the correction is discussed in detail in Appendix C.

\textsuperscript{3}Vickrey (1969) actually analyzed the case where the desired arrival times were spread equally over a peak hour.
Chapter 4  Modeling Combined Pricing and Information Systems

No-Toll Equilibrium: Single Route

This is the case where congestion pricing is not applied. Assume there are \( N \) commuters who wish to travel from their home to work along a single route. If \( D(t) \) is the length of the queue at time \( t \), then

\[
T_v^* = \frac{D(t)}{s}
\]  

(4.2)

Let \( T \) denote the most recent time at which there was no queue, and \( r(t) \) denote the departure rate function. Then

\[
D(t) = \int_{T - T_f}^{t - T_f} r(u) \, du - s(t - T)
\]  

(4.3)

and

\[
D'(t) = \frac{\partial D(t)}{\partial t} = r(t) - s
\]  

(4.4)

Equilibrium implies that travel costs are equal for all commuters, and that there is no way unilaterally to improve travel costs. The equilibrium departure rate \( r(t) \) is given by

\[
\begin{align*}
    r(t) &= s + \frac{\beta s}{\alpha - \beta} \quad t \in [t_q, \tilde{t}] \\
    &= s - \frac{\gamma s}{\alpha + \gamma} \quad t \in [\tilde{t}, t_q']
\end{align*}
\]  

(4.5)

(4.6)

where \( t_q, t_q' \) are the times at which the departures begin and end, respectively, and \( \tilde{t} \) is the departure time that results in on-time arrival, i.e.

\[
\tilde{t} + T_v(\tilde{t}) + T_f = t^*
\]  

(4.7)

Thus there are different constant departure rates for early and late departures (i.e. departures that result in early or late arrivals). The results are based on the requirement that in equilibrium, the marginal benefit from postponing departure by a unit of time equals the marginal cost of doing so. Thus for early departures the marginal benefit from postponing departure is the reduction in schedule delay costs, \( \beta(1 + D'(t)/s) \), which must equal the marginal cost of postponing due to the increase in travel time cost given by \( \alpha D'(t)/s \). Combining this condition with Equation 4.4 yields Equation 4.5. A similar approach gives Equation 4.6. The parameters that bound the time intervals are given by

\[
t_q = t^* - \left( \frac{\gamma}{\beta + \gamma} \right) \left( \frac{N}{s} \right) - T_f
\]  

(4.8)
4.4 Modeling Default Behavior using an Equilibrium Model

\[
t_q' = t^* + \left( \frac{\beta}{\beta + \gamma} \right) \left( \frac{N}{s} \right) - T_i \tag{4.9}
\]

\[
i = t^* - \left( \frac{\beta \gamma}{\alpha (\beta + \gamma)} \right) \left( \frac{N}{s} \right) - T_i \tag{4.10}
\]

Using these results, and bearing in mind that the first commuter to leave at \( t_q \) incurs no congestion\(^4\), his equilibrium travel cost \( C_{Eq} \) is given by

\[
C_{Eq} = \alpha T_i + \beta (t^* - t_q - T_i) \tag{4.11}
\]

\[
= \alpha T_i + \frac{\beta \gamma}{\beta + \gamma} \frac{N}{s} \tag{4.12}
\]

No-Toll Equilibrium: Two Parallel Routes

With parallel routes individual routes must of course still be in equilibrium (such that all commuters on a given route are in equilibrium); in addition, equilibrium across routes must be ensured, by making the costs on the two routes equal. Thus the equilibrium route split must be determined. Assuming that the total number of commuters on both routes is \( N \), the equilibrium route splits \( N_i, i = 1, 2 \) are given by

\[
N_i = \frac{s_1 s_2}{s_1 + s_2} \left[ \frac{N}{s_j} + \frac{\alpha}{\delta} (T_j^f - T_i^f) \right] \quad i, 1, 2 \quad j = 1, 2 \quad i \neq j \tag{4.13}
\]

In that case, equilibrium costs are given by

\[
C_{Eq} = \alpha \Gamma + \delta \left[ \frac{N}{s_1 + s_2} \right] \tag{4.14}
\]

where

\[
\Gamma = \frac{s_1 T_1^f + s_2 T_2^f}{s_1 + s_2} \tag{4.15}
\]

and

\[
\delta = \frac{\beta \gamma}{\beta + \gamma} \tag{4.16}
\]

\(^4\)Actually the first commuter should incur a time in passing through the queue equal to the unit service time, but this is ignored since we are dealing with a model that assumes "continuous" commuters, and the "first commuter" is thus infinitesimally small, similar to an infinitesimal flow of liquid in a pipe.
The Optimal Step-toll Equilibrium: Single Route

With a single step-toll in place, commuters pay a toll \( \tau \) when they pass the toll station between the period \([t^+, t^-]\)^5; at other times during the period of departures, no toll is charged. We are considering two cases for the toll station location: Upstream (of the bottleneck), at \( B \), and downstream at \( C \), because we wish to examine the impact of toll station location relative to the bottleneck on travel costs when incidents are assumed to occur. The derivation of the step-toll equilibrium and the optimal step-toll (the set \( \tau, t^+, t^- \) for which benefits [improvements in total traveler costs excluding tolls] are maximized) is provided in Appendix B for both cases.

An important feature of the optimal step-toll equilibrium is that the queue falls to zero at the beginning of the toll-period and at its end, but is never zero elsewhere within the period of departures. For equilibrium to hold, in the presence of any step-toll, there are no departures for a period \( \tau/\alpha \) during the travel period \( t_q, t'_q \). This “gap” in departures begins at that point in time which results in arriving at the toll station just as the toll period begins, denoted by \( t_\tau \). Since the last person to depart before the toll period begins must have the same cost as the first person who departs after the toll period begins, but the queue is never idle, they occur approximately the same schedule delay (differing only by a service time unit). But since the former does not pay the toll, \( \tau \), the latter compensates by leaving \( \tau/\alpha \) later and facing travel costs that are reduced by \( \alpha \cdot \frac{\tau}{\alpha} = \tau \), since schedule delay costs are (almost) identical. The other salient feature of the step-toll equilibrium is that there is a bulk departure of \( 2s\tau/(\alpha + \gamma) \) commuters just after the end of the toll period. If we assume that the position is random in the bulk (i.e. it averages out over a number of days) then the travel cost due to travel time and schedule delay for each commuter in the bulk must be higher than those of the last commuter to leave at the toll period ends. This is achieved by a bulk size of \( 2s\tau/(\alpha + \gamma) \). Aside from the gap and the bulk which are necessary for equilibrium when the step-toll is in place, the departure rates are unchanged from the no-toll case, since the marginal costs and benefits of postponing a trip are only a function of the capacity and the traveler’s characteristics. Thus the departure rates are

---

^Arnott et al. (1990a) argue that it is natural to express the toll as a function of arrival times (which obtains in their case because the travel time between the bottleneck exit and the destination is zero) “since toll gates are at the front of queues”. This is a fallacious argument, not because toll gates can or should not be located at the front of points of congestion, but because the tollgates referred to are those that currently exist to collect (time-invariant) revenue generating tolls, and the congestion is actually caused by the toll gate. See Seila and Wilson (1991) for a discussion of the efficiency issues arising when the toll collection process contributes to the congestion problem! These problems are diminishing in significance due to the increased use of electronic toll collection systems which charge travelers as they drive through the “toll booth” at highway speeds. On the other hand, it is the opinion of at least some turnpike managers that the old-fashioned toll stations serve as a useful metering device that prevents or mitigates downstream congestion. But here the argument would be that with congestion pricing, this device would be unnecessary due to reduced congestion.
4.4 Modeling Default Behavior using an Equilibrium Model

given by

\[
\begin{align*}
  r(t) &= s + \frac{\beta s}{\alpha - \beta} & \quad t \in [t_q, t_r) & \quad (4.17) \\
  &= 0 & \quad t \in (t_r, \tilde{t}) & \quad (4.18) \\
  &= s - \frac{\gamma s}{\alpha + \gamma} & \quad t \in (\tilde{t}, t_q') & \quad (4.19) \\
  &= 2s\tau/(\alpha + \gamma) & \quad t = t_q' & \quad (4.20)
\end{align*}
\]

where the notation is as for the no-toll case. The values of the parameters for the optimal step-toll are given below. The optimal toll minimizes travel costs excluding tolls and is given by

\[
  \tau = \frac{\delta N}{2s} \tag{4.21}
\]

The first departure occurs at

\[
  t_q = t^* - T^f - \frac{\gamma}{\beta + \gamma} \left( \frac{N}{s} \right) + \frac{(\gamma - \alpha)\tau}{(\beta + \gamma)(\alpha + \gamma)} \tag{4.22}
\]

The toll period begins, for the upstream case, at

\[
  t^+ = t_q + \frac{\tau(\alpha - \beta)}{\alpha\beta} + T^f - T_{TS}^f \tag{4.23}
\]

where \( T_{TS}^f \) is the (constant) travel time from the toll station location to the entrance of the bottleneck at the end of the route.

For the downstream case

\[
  t^+ = t_q + \frac{\tau}{\beta} + T^f \tag{4.24}
\]

Since the last vehicle faces no queue, the toll period ends for the downstream case at

\[
  t^- = t_q + \frac{N}{s} - \frac{2\tau}{\alpha + \gamma} + T^f \tag{4.25}
\]

and for the upstream case at

\[
  t^- = t_q + \frac{N}{s} - \frac{2\tau}{\alpha + \gamma} + T^f - T_{TS}^f \tag{4.26}
\]

The equilibrium travel cost \( C_{Eq} \) for each traveler is given by

\[
  C_{Eq} = \alpha T^f + \delta \left[ \frac{A N^*}{2s} - 2\Delta \right] \tag{4.27}
\]

where \( A = \frac{3(\beta + \gamma)(\alpha + \gamma) - \beta(\gamma - \alpha)}{2(\beta + \gamma)(\alpha + \gamma)} \).

The total net travel cost on a route (excluding toll revenues) for the optimal step-toll is given by

\[
  TC = \alpha T^f N + \frac{\delta AN^2}{2s} \tag{4.28}
\]
Chapter 4  Modeling Combined Pricing and Information Systems

The Optimal Step-toll Equilibrium: Two Parallel Routes

For two routes in parallel, we need to find the route split \( N_1, N_2 \) which minimizes total travel costs excluding tolls, assuming that the optimal step-toll (for a given \( N_i, [i = 1, 2] \)) is in place on each route. At the optimal route split, however, the equilibrium costs on one route may not equal those on the other, since only travelers on a given route are of necessity in equilibrium. If they differ, equilibrium across routes may be achieved by applying a uniform toll \( \tau_s \) on the lower cost route (which is incurred at all times when traveling on that route), in addition to the step toll. This does not affect the optimality of the solution because the uniform toll consists entirely of toll revenue, which does not add to the total cost to be minimized. The uniform toll necessary to obtain the optimal route split was incorrectly derived by Arnott et al (1990b). The correction is detailed in Appendix C. The correct value is given by

\[
N_i = \frac{s_1 s_2}{s_1 + s_2} \left[ \frac{N}{s_j} + \frac{\alpha}{\delta A} (T_j^f - T_i^f) \right] \quad i = 1, 2 \quad j = 1, 2 \quad i \neq j
\] (4.29)

The value of the uniform toll that results in the optimal route split is given by

\[
\tau_s = \frac{\alpha (\beta + \gamma)(\alpha + \gamma)}{3(\beta + \gamma)(\alpha + \gamma) - \beta(\gamma - \alpha)} (T_2^f - T_1^f)
\] (4.30)

assuming that \( T_2^f > T_1^f > 0 \).

Equilibrium costs for that case are given by

\[
C_{Eq} = \alpha T_1^f + B \delta \frac{N_1}{s_1} + \tau_s = \alpha T_2^f + B \delta \frac{N_2}{s_2}
\] (4.31)

where \( B = \frac{2(\beta + \gamma)(\alpha + \gamma) - \beta(\gamma - \alpha)}{2(\beta + \gamma)(\alpha + \gamma)} \) (see Appendix C).

4.5 Modeling Incidents

The most important characteristic of (most) real-world incidents, such as accidents or breakdowns, is their randomness. Certain elements of the automobile/highway infrastructure entity can be improved or modified to make incidents less likely or less severe. For example, automobile braking systems are continually refined to make accidents less likely, and engineering improvements make cars more reliable and hence less likely to break down. Intersections with high accident rates are often redesigned for higher safety, sometimes physically restructuring them, such as the replacement of a regular intersection with an at-grade separated one. But the occurrence of incidents remains fundamentally unpredictable. Thus one simply does not know in advance the location, severity (in terms of reduction of capacity of affected link), and duration of traffic incidents. Moreover the temporal structure of the
capacity reduction can vary depending, for example, on the response pattern of emergency vehicles summoned to "remove" an accident.

Giuliano (1989) has investigated the characteristics of such incidents by analyzing data from a major Los Angeles freeway. Results show that accidents make up a very small proportion of all incidents, but account for a greater share of total incident duration.

We will model incidents in such a way as to capture this basic randomness, but maintain simplicity. We will evaluate the performance of the system in the presence of incidents by limiting our analysis to the case of a single incident. Thus a conditional analysis is performed that looks at the effects of a single incident, since we believe this to be representative of the general impact of incidents - a random, limited time reduction in network capacity. We make no assumptions about the frequency of occurrence of such one-incident days - an empirical question that relates to the particular network characteristics. Thus the network may experience days on which two or more incidents occur, and days which are incident-free. We have implicitly assumed that incidents occur not too frequently because default behavior is based on incident free capacities, but this could possibly be overcome by replacing these capacities with some form of expected values that take the incidents into account. Furthermore, the impact of the incident on capacity of the link is modeled as a constant reduction in capacity of the affected link. Although this is unlikely to be the case in practice, it is simpler; moreover we are not aware of research that has detailed the form of the capacity reduction during an incident.

The reduction factor $F$, the time of occurrence $T_s$, the duration $D$, and the location of the incident are all assumed uniformly distributed random variables. Specifically, an incident is assumed equally likely to occur on either route. An incident will reduce the capacity of one of the two bottlenecks by a factor $F \sim U(0.35, 0.65)$. An incident is assumed to start within approximately one hour of the first arrivals of the bottleneck, so that the time of occurrence in minutes after the bottlenecks are busy is given by $T_s \sim (0, 60)$. The incident lasts for a duration in minutes $D \sim (15, 35)$.

### 4.6 Modeling the Response to Incident Congestion

In response to information about incidents, adjustments to default departure times and / or routes are made. We first describe how the information system operates, followed by the assumed user response.

#### 4.6.1 The Route Guidance System

En-route information is modeled as a guidance system that operates at node $B$. We assume that it is a perfect system in the sense that it "knows" (i.e. can predict perfectly) the end of
an incident after it begins and because there is no delay between the incident's occurrence and the provision of information (i.e. we assume the incident reporting delay is zero). Also, information is given to all travelers, and is based on minimum predicted cost.

When an incident occurs on link 1 or 2, the information system begins operation. At that point, the number of vehicles on both of these links, and the queue lengths, are obtained from the simulation program. The information system is also informed about the magnitude, duration, and location of the incident. Thus, the exit time of the last vehicle on each link to have passed B before the incident can be computed. Since we know the duration of the incident, we can accurately predict the exit time for a vehicle arriving at B on each link, in the following manner: The arrival time of the vehicle plus the free-flow travel time on the link result in the arrival time at the queue. If the last vehicle to exit is still in the queue, then the predicted exit time for the vehicle now at B is the exit time of the last vehicle plus the service time of the queue. Otherwise, we simply add the service time of the queue to the arrival time at the queue. The time of the last vehicle to exit each link is then updated according to the route choice made. If the predicted cost of travel on the default route is greater than that of the alternate route by more than a small threshold (to avoid excessive sensitivity), the vehicle is directed to that route. When analyzing the priced regime, the information is provided on the basis of the total cost including the tolls.

4.6.2 The Response to Guidance

All travelers are assumed to follow the guidance directives, which is in line with our idealized assumptions about a perfect information system. An issue of great importance in practical applications is the actual compliance rate, which depends on the number of vehicles that receive information, and the likelihood that they will comply. Kaysi (1992) observed during simulation experiments that the adverse impacts of information provision (concentration and overreaction) increased for higher compliance rates. Benefits due to information were maximized at compliance rates of 0.2 and 0.3. In his analysis the maximum temporal update frequency (corresponding to the minimum time interval at which the information system updates itself) was 2.5 min. Since in our case the information system is updated continuously, overreaction and concentration are eliminated; as a result, benefits will likely be overestimated. It is still possible that higher benefits in terms of total system costs could be obtained with lower guided probabilities, since our system, in spite of its perfectness, does not attempt to minimize system costs. In the analysis of the next chapter, lower guidance rates will be examined for the case of the unpriced regime to investigate this point, and for the case of the tolled regime in order to overcome a problem that was observed with 100% guidance.

We shall be carrying out a conditional analysis which examines the case of only a
single incident, and are therefore at least theoretically in a position to compute the system-optimal travel pattern, given that the incident has just occurred. As previously discussed, however, we think that the system optimal approach is of less interest due to the difficulty of computing the guidance and due to its potential implementability problems.

### 4.6.3 Simulation Model

A macroscopic simulation model developed at MIT by Mukundan (1992) was adopted to evaluate the network performance. Travel time on a link is the sum of the time spent on the "moving" part, whose speed-density relationship can be specified by the user; and the time spent in the queuing part at the end of the link, where users may have to queue before entering the next link. The simulator is thus perfectly suited for modeling the bottleneck model we are using. The default route and departure time patterns were calculated using the equilibrium model, and discretized for use in the simulation. The simulator was modified to incorporate the information system. The information system can provide minimum cost guidance to travelers that depend on the regime. In other words when congestion pricing is applied, it provides certain summary statistics for each traffic cell, which we have further processed to yield the desired statistics. Note that simulators such as this model discrete vehicles. Since equilibrium models (including the one used here) generally are based on real-valued representation of commuters, the simulation can only approximately the analytical equilibrium flow pattern obtained by these models. The impact of these approximations is discussed somewhat further in the next chapter, when specific references are made to the model used.
Chapter 5

Combined Pricing and Information Systems: Analysis

This chapter explores several of the integrated design issues described in the previous chapter in a quantitative setting. These include toll station location, information provision to the bulk, superadditivity, the impact of forcing the toll structures to be identical on different routes, and the consequences of providing a "free" alternative by tolling only route.

We begin with some remarks about the impact of using simulation to approximate the equilibrium model. Then we discuss the results obtained when simulating the effects of information supplied to both no-toll and toll regimes for the network described in the previous chapter.

The general pattern of the benefits arising when information is supplied to a no-toll regime is then discussed. This includes a description and interpretation of the cost curves and the effect of lowering the guidance rate.

The impact of information on a toll regime is then analyzed. We look at two cases of toll regimes, one in which the toll stations are located downstream of the bottleneck, and the other in which they are situated at the entrance to the link (upstream of the bottleneck), and discuss the implications when incidents and information are introduced. The effect of modifying the information system is explored, including the case in which information is not supplied to the bulks at the ends of the toll periods. Toll structures that are identical on the two routes, and easily comprehensible to the user, should result in increased implementability. Deriving optimal tolls under these conditions, and examining their impacts, is discussed subsequently. Next, optimal tolls are derived when only one route is tolled, and their interaction with incidents and information assessed. The question of whether there are synergistic effects when applying pricing and information jointly is then discussed. Conclusions are drawn at the end of the chapter.
5.1 A Note on Simulation Results

We assume in the analysis described in this chapter that the desired arrival time at $C$, $t^*$ = 8.0 hrs, i.e. 8:00 a.m. However, for greater clarity and ease of visualization, we shall generally use times in simulation seconds, i.e. the number of seconds after the first vehicle departs (the earliest time chosen for departures on the tolled and untolled regimes). Based on this transformation, $t^*$ corresponds to a time in seconds of 7622. The free-flow travel times on links 3, 1, and 2, are 900, 720, and 1080, respectively, and hence the free-flow travel times on routes 1 and 2 are 1620 and 1980, respectively (refer to Figure 4-1). Unless otherwise noted, the x-axis of the figure denotes the arrival time at the “information node” $B$, in simulation seconds. This will make it clear whether the vehicle was in a position to receive information, based on the incident start times.

We present the equilibrium costs for both the untolled and the step-toll case as a function of the departure times to illustrate the consequences of the approximation due to using a simulation model, and to show the variation in the costs experienced by the travelers leaving in the bulk just after the toll ends.

Figure 5-1 plots the departure time in simulation seconds from the origin (at $A$) against the total travel costs for the entire set of commuters when we have no tolls and when no incident occurs. Ideally this should be a straight line, and hence this is obviously not an equilibrium in the strict sense. Aside from the few outliers, which arise because of a small error in the simulation model, the observed variations in total cost are simply due to the fact that we approximate a continuous function (the equilibrium model described in the previous chapter) by a discrete (in time and travelers) one (the MIT simulator). Note that the pattern displays greater variability after the departure which results in on-time arrival, since the departure rates decline substantially after that point, which generates larger errors in the approximation, and because the value of time late is greater than that of time early, which means a given approximation in time results in greater cost penalty. Lower departure rates imply a coarser approximation to the analytical solution, which results in higher deviations from the “true” values. The discretization itself must be done with care. Since the simulator’s default way of generating departures consists of taking the number of travelers known to depart in a standard time interval (we employed one-minute intervals) and assigning random departure times within that period to each. Using that method we obtained badly skewed costs. Only by assigning departures that are uniformly distributed, and begin after the first departure time calculated by the analytical model, were travel costs obtained that are both unbiased and of low variance.

Figures 5-2 and 5-3 show the equilibrium cost pattern when pricing is applied. In Figure 5-2, the total travel cost does not include any tolls that may have been paid. Due to the larger scale used on the vertical axis, the minor variations in travel cost (before the
spikes at the right of the figure) appear more smooth. Note that the travelers who depart after the gap experience a decrease in travel costs since the queues have diminished as a result of the gap. The bulk departure at the end of the toll period on each route is such that the “first” commuter in the bulk experiences very little cost. Those who leave subsequently experience growing costs due to the queue ahead of them. It is clear from the diagram that the commuters in the bulk are not in equilibrium on a given day. As discussed previously, the assumption is that the positions in the queue on a given day are equally likely, and thus the commuters are, “on average”, in equilibrium. For the purpose of this chapter we also make the additional assumption that, despite the presence of the information system, the non-equilibrium pattern in the bulk will not induce switching if no incident is in progress.

Figure 5-3 shows the equilibrium costs including tolls when pricing is applied. It can be seen that the travelers prior to the bulk are in approximate equilibrium, and that the cost of the travelers in the bulk is, on average, equal to that equilibrium cost. Instead of all departing at one point in time, the simulation model spreads the commuters in the bulk on each route over a small period in time (1 minute).

Applying the results of the previous chapter both the no-toll and the step-toll equilibrium to our network parameters yields the results displayed in Table 5.1. The uniform toll is applied to route 1, which has the shorter free-flow travel time, and equals $0.22.
Figure 5-2: Equilibrium costs for toll regime (excl. tolls)

Figure 5-3: Equilibrium costs for toll regime (incl. tolls)
5.2 Incidents and Information in the No-Toll Regime

<table>
<thead>
<tr>
<th>Regime &amp; Route</th>
<th>N</th>
<th>EC</th>
<th>Toll($)</th>
<th>T⁺</th>
<th>T⁻</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untolled 1</td>
<td>1634</td>
<td>9.38</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Untolled 2</td>
<td>566</td>
<td>9.38</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Tolled 1 Upstream</td>
<td>1620</td>
<td>9.28</td>
<td>3.00</td>
<td>6:33</td>
<td>7:59</td>
</tr>
<tr>
<td>Tolled 2 Upstream</td>
<td>580</td>
<td>9.28</td>
<td>3.22</td>
<td>6:32</td>
<td>7:53</td>
</tr>
<tr>
<td>Tolled 1 Downstream</td>
<td>1620</td>
<td>9.28</td>
<td>3.00</td>
<td>7:16</td>
<td>8:12</td>
</tr>
<tr>
<td>Tolled 2 Downstream</td>
<td>580</td>
<td>9.28</td>
<td>3.22</td>
<td>7:18</td>
<td>8:11</td>
</tr>
</tbody>
</table>

Table 5.1: No-Toll equilibrium and equilibrium with optimal step-toll

5.2 Incidents and Information in the No-Toll Regime

In this section we describe the results obtained when providing information (in our case, route guidance at node B) when incidents occur in the network. A single incident is assumed to occur either on link 1 or 2, and is assumed to affect the capacity of the respective bottlenecks for simplicity. (The free-flow travel time could alternatively be assumed to increase as a result of an incident). One of our main objectives in the next section is to evaluate the impact of toll station location on various performance measures. Since the toll stations are placed either at the beginning (at B) or end (at C) of links 1 and 2, our basic objective of modeling random increases in travel time on these links is served adequately by restricting the impact of incidents to capacity reductions in the bottlenecks. This restriction on the location of the incident is somewhat unrealistic because in practice incidents may occur anywhere along a highway, but should not affect the results beyond making the accidents more severe than if their location were allowed to vary (since for a given magnitude and duration of incident, more people are affected if the location is at the bottleneck than upstream of it).

Recall that our information system is “perfect” in the sense that it “knows” both when an incident ends, and the magnitude of the capacity reduction. The capacity is assumed to remain constant at its reduced level throughout the incident’s progress. The variables used in the title bar of each figure are \( t_a, t_e \), for start and end time of incident in simulation seconds, \( loc \) for the link on which it occurred, and \( fac \) for the capacity reduction factor due to the incident, i.e. the capacity of the affected bottleneck during the incident is given by the product of \( fac \) and the default capacity of the bottleneck. The legends R1, R2 refer to travelers with default routes 1 and 2 respectively, R1-R2 refers to commuters with default route 1 switching to route 2, and R2-R1 to commuters switching from 2 to 1.
Chapter 5  Combined Pricing and Information Systems: Analysis

Figure 5-4: Costs with incident on route 1 - Untolled regime, No Information

5.2.1 Cost Curves

Figures 5-4 and 5-5 show the travel costs as a function of arrival time at B for an incident on route 1 for the case of no information and information supplied, respectively. The legend “R1->R2” denotes a traveler with default route 1 who was switched to route 2.

Examining 5-4, it is seen that the shape of the cost curve for the commuters on route 1 has a number of kinks, which we will now explain. (The depicted cost curve for route 2 is of course flat, since it is unaffected by the incident in the absence of information). The costs first begin to rise at the moment the incident occurs. As long as the incident is in progress, delays are increasing, and hence the curve climbs during this phase. Once the incident is over and the capacity returns to its normal level, there is a constant delay imposed on all commuters who exit the system after this point due to the additional travel time incurred by all those who passed the bottleneck at reduced capacity. This delay equals the product of the duration of the incident and the ratio of the reduction in service rate to the original service rate. However, since costs are a combination of travel costs and schedule delay costs, the costs are not uniformly higher after the end of the incident. The commuters who arrive before $t^*$ when there is no incident and still arrive before $t^*$ after the incident are responsible for the (short) flat part of the curve following the initial rise, because their travel costs are uniformly increased by the product of the delay and the difference in unit costs of travel.
5.2 Incidents and Information in the No-Toll Regime

Figure 5-5: Costs with incident on route 1 - Untolled Regime, With Information

Figure 5-6: Costs with incident on route 2 - Untolled Regime, No Information
time and schedule delay (i.e. delay times $|\alpha - \beta|$).

This is followed by a period of increasing costs, which is borne by those commuters who as a result of the incident arrive after $t^*$ whereas they would arrive before $t^*$ if there were no incident. As a result, the (constant) delay is responsible for increasing schedule delay costs. Finally, the curve is flat till the end in that phase during which all the commuters would arrive after $t^*$ even without the incident, and thus experience a common increase in costs (equal to $|\alpha + \gamma|$ times the common delay due to the incident).

In 5-5, the same incident is depicted, with the information system in place. Before the occurrence of the incident, the cost curve is identical to that of the no-information case. As soon as the incident begins, the vehicles at, or upstream of, $B$ will receive route guidance. What follows is a period during which vehicles are shifted from route 1 onto route 2, which is operating at default capacity. This period ends when the shifting has increased the expected costs on route 2 and decreased them on route 1 to the point that they are equal (or differ by the specified switching threshold). For the following phase the cost curve is flat as switching has stopped. Now, our network parameters are set so that the last departure on route 2 occurs some time before departures end on route 1. Thus, during the period following the last departure on route 2, commuters are switched from route 1 to route 2 as deemed necessary by the information system, since the costs on route 2 will decrease when
Figure 5-8: Benefitogram - Incident on route 1, Untolled Regime

Figure 5-9: Benefitogram - Incident on route 2, Untolled Regime
departures cease\(^1\). As a result, this last batch of commuters experiences declining costs, which is seen by the (short) dipping part of the cost curve. Figures 5-6 and 5-7 show the same situations when the incident is on route 2, and have the same basic structure.

**Benefitograms**

Another way to look at the impact of information on travel costs is through the “benefitogram”, which essentially combines the cost curves for the cases with and without information for a particular incident. In a benefitogram we show cumulative benefits as a function of departure time, by subtracting the costs a particular commuter incurs in the presence of information from those incurred when no information is supplied. The construction of benefitograms is facilitated by the deterministic departure rates that we have assumed. We show the benefitograms for the two examples just presented in Figures 5-8 and 5-9. The general structure is as follows: A period of (declining) growth is followed by a leveling off of benefits, followed by a rise or dip depending on the location of the incident. As explained above, this last part of the curve is made up of route 1 commuters, hence, if the incident is on the other route they will all be worse off (due to the shifting of route 2 commuters on their route) compared to the no-information case. If the incident is on route 1 they will all be better off, since some of the earlier commuters on the route have shifted to route 2 and thus queue lengths are shorter.

**5.2.2 Reducing the Guidance Rate**

We also tested the impact of lower guidance rate on improvements in system costs. Figure 5-10 shows, for 100 runs, the average improvements obtained with guidance rates of 100%, 50%, 25%, and 10%. The improvements obtained with a 50% guidance rate were almost as high as when 100% guidance was used. 25% and 10% rates resulted in significantly lower benefits.

The benefitogram climbs more slowly to its peak value with lower guidance rates. In Figure 5-11 we show the benefitogram for the same incident on route 1 as before but with the guidance rate at 25%. These results contrast with those of Mahmassani and Chang (1991) and Kaysi (1992), who reported highest benefits at low guidance rates. This is because the information systems were not perfect like the one assumed here. With a perfect information system, higher guidance rates do not necessarily result in the ill effects of providing information, such as overreaction.

\(^1\)Note that in the no-incident case, the batch of route 1 commuters is not shifted to route 2 after departures stop on route 2 due to equilibrium.
5.2 Incidents and Information in the No-Toll Regime

Figure 5-10: Average Improvements for various guidance rates

\[ ta = 4740 \quad te = 5940 \quad loc = 1 \quad fac = 0.521 \]

Figure 5-11: Benefitogram - Incident on route 1, 25% Info, Untolled Regime
5.2.3 The Magnitude of the Benefits

The magnitude of the average reductions in total system cost (sum of total travel costs over all commuters) due to information are shown for a 100 run experiment in Figure 5-12. These reductions are "small", for the following reasons:

- Links 2 and 3 are relatively long. Therefore, when the incident occurs, a rather large number of people who are on the affected link are stuck with no ability to adjust their routes, and thus the information benefits fewer people.

- Because of the long free-flow travel times "fixed" travel costs constitute a large fraction of total travel costs.

- Capacity is not added by the information system; that is, the bottleneck capacities are not affected by the provision of guidance. (We will address this point further in Appendix D).

The magnitude of the benefits obtained here is generally consistent with results obtained by other researchers. An example is the work of Koutsopoulos and Lotan (1990), which reported gains of 4.4% in total cost. The information system used in their research differed from ours in that it was used to reduce the error in perceived travel times, not to reroute as a result of incidents. A similar modeling approach to the one employed here was used by Kaysi (1992), who reported maximum reduction in total cost of 4.3%. The network employed in that work consisted only of bottlenecks. This contrasts with our network, in which fixed travel costs make up a large share of total travel costs. Hence improvements should be less in our case, ceteris paribus.

On the other hand even these small improvements could be overstating the benefits due to information. Our information system was assumed to be perfect, and the figures just discussed refer to the case where everyone both had access to information and complied with the directives. In addition we make the assumption that in the absence of information travelers strictly adhere to their default behavior. This is of course somewhat unrealistic since nowadays radio traffic reports are ubiquitous, and even without them many incidents can either be seen directly or perceived through excessive delays, prompting travelers to make route adjustments.

5.3 Incidents and Information in the Toll Regime

In this section, we will investigate the impact of information on a priced regime. This section addresses some of the issues arising in the joint implementation of congestion pricing and traveler information systems. Thus, the optimal step-tolls computed in the previous chapter
5.3 Incidents and Information in the Toll Regime

![Histogram of benefits - Untolled regime](image)

Figure 5.12: Histogram of benefits - Untolled regime

are assumed to be in place, and the default travel patterns are assumed to be based on them. The primary design issue investigated in this section is the toll station location. In addition we examine the provision of information to the commuters who constitute of the departure bulks at the end of the toll period on each route. Two basic situations will be examined: One in which the toll stations are upstream of the bottleneck (in the case of our network, at B), and one in which they are located between the bottleneck and the destination (at C). Of course, without incidents, the choice of location is immaterial for our model, since the toll periods can always be adjusted appropriately so that the toll costs as a function of departure time are identical.

First the two cases are studied when the information system is in its base configuration, guiding all commuters over the entire period following the incident. Modifications to the information systems that are expected to improve the performance are then explored.

5.3.1 Base Case Results

Our analysis is similar to that of the previous section in that we both analyze particular incidents to obtain a detailed look at the workings of the combined system and the aggregate performance over a large number of runs for observing general trends.

We will look at cost curves and benefitograms both when toll costs are included and
when they are excluded. Since the latter case represents only a component of the total costs it might not be meaningful to look at the benefitograms since there is no reason one would expect any particular structure, given that the information system operates on total cost. But since it comprises (sunk) travel cost, which is what we are trying to minimize through imposition of the toll, we think it is useful to consider its behavior.

**Upstream Toll Stations**

Figure 5-13 shows the cost curve (excluding tolls) when there is an incident on route 1, with the toll stations at B, and no information. It differs in general from the no-toll cost curves only in the departure gap, and the bulk phenonema at the end of departures on both routes. Figure 5-14 shows the same situation but with information supplied.

The corresponding benefitogram is shown in Figure 5-15. Note that it differs in shape from the no-toll benefitograms in two respects. First, there is a small drop in the “benefits” immediately after the start of the incident (note that these are not really benefits from the user’s perspective as tolls are excluded). Prior to the incident, the routes are in equilibrium. Therefore the higher tolls on route 1 imply that travel costs excluding tolls are lower on that route than on route 2, as long as the toll periods are “on” on both routes. After the occurrence of the incident, costs on route 1 (without info) begin to rise. (Recall that we are subtracting the costs with information from those without to get benefits). During a short period after the incident begins, the costs have have not grown sufficiently on route 1 to outweigh the initial “advantage” over route 2, making things comparatively worse for those switched during that phase. In addition, the benefitogram also has a U-shaped portion at the end. (It has an inverted U-shape when incidents occur on route 2). The explanation is similar to that for the shape of the kink at the end of the benefitogram for the case of the untolled regime, except that the sloping shape is replaced by a “U” since the last departures primarily consist of the two bulks, which are large number of commuters departing over a very small time interval (the portion of the benefitogram that connects the two spikes is generated by the commuters on the route whose bulk departs later). Thus the positive spike is generated by the travelers whose default route suffered the incident, and the negative part by those whose default route was incident-free, since, in general, the former are made better off as a result of information and the latter, worse.

Similarly, Figures 5-16 and 5-17 show the cost curves with and without information when costs include tolls, and Figure 5-18 shows the corresponding benefitogram. The benefitogram has an interesting attribute in the form of a little dip after benefits have begun to climb. It can be explained as follows. Commuters are no longer being redirected since costs are equal on the two routes, $t_2 < t_1$. Thus, costs will rise disproportionately on route 2, since the late arrival factor $\gamma$ is greater than $\beta$, and switching away from 2
5.3 Incidents and Information in the Toll Regime

will begin to counteract this trend, until costs are equal. A similar event occurs just after \( t_1 \), resulting in a period of switching from 1 to 2. Switching vehicles from the no-incident route to the incident route produces negative contributions to the benefitogram. This arises because the cost on the incident route is always greater than the cost experienced by the switched vehicle in the absence of information on the incident-free route, and hence the first of the two brief switching phases just described accounts for the drop in benefits. The general structure of the curves is similar when the incident is on route 2.

General observations in regard to upstream toll station locations include the following.

1. In the presence of incidents, and without information, toll costs are not affected by the incident, since the delays due to the incident arise due to queuing in the bottleneck, which is located after (downstream of) the toll stations. This holds for the particular incident examined here, as seen in Table 5.2, where the number of toll paying travelers is the same for the (no-incident) equilibrium case and the no-information incident case.

2. Information always produced improvements in total cost excluding tolls. The net contribution of the U-shapes at the end of the benefitograms was generally positive when costs excluded tolls. In contrast, the contribution of that part to the benefitogram for costs including tolls was, on some occasions, sufficiently negative to offset the entire benefits previously accumulated; hence providing information may result in negative net benefits. In a sense, therefore, the bulk departures are "responsible" for the possible increases in total cost when toll costs are included in the upstream case. Of course, since we are using a myopic system, we cannot really make strong statements in this regard because the costs on the two routes faced by the bulk are a result of what has occurred in terms of guidance before that point in time. Nevertheless we feel inclined to attribute this phenomenon to the discontinuous nature of the step-toll and the non-equilibrium cost pattern of the bulks, which are in equilibrium only in an average sense. We investigated this phenomenon further by plotting the benefits against the start time of the incident when the incident was on link 1 in Figure 5-19 and when it was on link 2 in Figure 5-20. Figure 5-19 shows little variation of the spread of benefits with start time of incident. Figure 5-20, however, indicates that beyond a certain time, benefits are always negative. Since total costs excluding tolls always decrease with information, the total toll revenue must have increased compared to the no-information case. Figures 5-21 and 5-22 show the total number of toll paying commuters when the incident is on route 1 and route 2, respectively. In the no-information case, 1003 travelers pay the step-toll on both routes. When the incident is on route 1, the switching between the bulks produces a net increase in the number of toll paying travelers of about 25. This is obtained when some travelers on route 1, whose toll period ends several minutes after that on route 2, switch to
Figure 5-13: Costs (excl. tolls) - Incident on Route 1, No Info, Upstream Toll Station

route 2 (reducing the number compared to the no-information case), followed by the switching to route 1 of a subset of the bulk on 2, which increases the number of toll paying travelers. When the incident is “late” on route 1, however, fewer travelers will switch from the bulk on route 2, which gives rise to the lower values observed in Figure 5-21 for high values of start time of incident. When the incident is on route 2, the opposite trend is observed, as shown in Figure 5-22. In particular, late incidents will increase the number of bulk travelers switching to route 1. Finally, whether the net benefits will be negative depends not only on the change in the number of toll paying travelers (or more precisely, the change in total revenue as the tolls differ somewhat on the two routes), but also on the benefits excluding tolls. This explains why for a given number of toll paying travelers, net benefits may be positive or negative.

**Downstream Toll Stations**

Figure 5-23 shows the cost curve excluding tolls when there is an incident on route 1, with the toll stations at C, after the respective bottlenecks, and no information. Figure 5-24 applies to the same situation with information supplied. The corresponding benefitogram is shown in Figure 5-25. Similarly, Figures 5-26 and 5-27 show the cost curves without and with information when toll costs are included, and the corresponding benefitogram is given.
5.3 Incidents and Information in the Toll Regime

Figure 5-14: Costs (excl. tolls) - Incident on Route 1, With Info, Upstream Toll Station

Figure 5-15: Benefitogram (excl. tolls) - Incident on Route 1, Upstream Toll Station
Figure 5-16: Costs (incl. tolls) - Incident on Route 1, No Info, Upstream Toll Station

Figure 5-17: Costs (incl. tolls) - Incident on Route 1, With Info, Upstream Toll Station
5.3 Incidents and Information in the Toll Regime

![Graph](image)

**Figure 5-18:** Benefitogram (incl. tolls) - Incident on Route 1, Upstream Toll Station

![Graph](image)

**Figure 5-19:** Benefits for Upstream Locations: Incident on Route 1
Figure 5-20: Benefits for Upstream Locations: Incident on Route 2

Figure 5-21: Number of Toll Paying Commuters with Information: Incident on 1
5.3 Incidents and Information in the Toll Regime

![Diagram of start time of incident vs. paying toll]

Figure 5-22: Number of Toll Paying Commuters with Information: Incident on 2

in Figure 5-28. These curves are somewhat more complex than those for the upstream case, as explained below.

Figure 5-24 reflects the fact that there is a single drop in travel costs (excl.tolls) when information is supplied. The benefitogram for the case including tolls (Figure 5-25) can be understood using Figures 5-23 and 5-24 from which it is derived. Thus, the up and down sloping portions of the benefitogram that occurs before the bulk departures are simply determined by whether the cost with information curve (Figure 5-24) is above or below the aggregate value of the costs without information depicted in Figure 5-23.

Figure 5-26, where no information is supplied, shows that a small number of travelers whose default arrival time at C is before the onset of the toll period are now paying the toll because of their delayed arrival at the toll station. In addition, the drop in cost due to the end of the toll period occurs “earlier” (i.e. for travelers who usually pay the toll), and thus a number of commuters will avoid the toll they usually pay. Figure 5-26 shows this phenomenon (in the brief portion of travelers with high costs before the gap). Figure 5-27 consists of three segments after the incident begins. The drops in cost that cause this fragmentation are the ends of the toll periods on the two routes.

The shape of the benefitogram (in Figure 5-28) can again be understood by noting that the costs with information (see Figure 5-27) are alternatively above and below the corresponding cost curve without information (see Figure 5-26). This is particularly easy
to see because the curve in Figure 5-26 is flat for the relevant portions (at a value of approximately $9.75), and thus the two successive situations where the benefitogram dips and then rises are directly related to being above or below that line.

Several important observations can be made in regard to downstream toll station locations in general.

1. **An incident may reduce the total number of toll paying travelers.** Incidents reduce the rate at which travelers exit the bottleneck. If the period during which the incident is in progress overlaps the toll period, fewer travelers pass through the toll station during the toll period. Total revenues decrease in this case. Thus a certain number of commuters who would pay a toll in the incident-free case now exit the bottleneck late enough to avoid the toll. As a result the costs including tolls without information were generally higher for upstream than downstream toll station locations, which can be seen from Figure 5-29, which plots the difference between upstream and downstream locations against the end time of the incident. For incidents that occur “early”, no difference exists between the two locations because the incident occurs before the start of the toll period. For our network the toll period at the downstream location begins at approximately 4900 and 5110 simulation seconds for links 2 and 3 respectively, and Figure 5-29 shows that the difference is constant at zero for incidents ending before 4900. When information was supplied the differences were somewhat greater, as shown in Figure 5-30. This is because information in the upstream case generally increases revenue, and increases costs compared to the downstream case, in which we observed no or negligible increase in the number of toll-paying travelers due to information.

2. On the other hand, the time shift in arrivals at the bottleneck caused by the incident may draw travelers into the toll period who do not usually pay the toll, resulting in significantly increased costs for them. This occurs when the incident begins before the start of the toll period, and may occur even when information is provided, since some travelers will be between B and C on the incident route when the incident occurs. If the incident ends before the toll period, the number of travelers passing through the toll station is unchanged compared to the no-incident case, since some travelers are drawn into the toll period who do not pay in the no-incident case and an equal number who usually pay arrive at the toll station after the toll period ends.

3. Similar to the upstream case, total travel costs excluding tolls always decreased with information. Moreover, total travel costs including tolls also invariably decreased with information. Note that since the toll station location is effectively just before the destination at C, equilibrium requires that the ends of the toll period are close to
5.3 Incidents and Information in the Toll Regime

\[
t_{a} = 4740 \quad t_{e} = 5940 \quad \text{loc} = 1 \quad \text{fac} = 0.521
\]

Figure 5-23: Costs (excl. tolls) - Incident on Route 1, No Info, Downstream Toll Station

each other, differing only sufficiently to make up for differences in the toll levels on each route, and hence there is little opportunity to switch to a route that is tolled from one that is not, or vice versa.

5.3.2 Modifying the Information System

In the previous section it was seen that, when the toll station is located upstream, the net contribution of the bulks to the benefitogram was negative. To address this problem, we will test the impact of two modifications to the information system: A reduced guidance rate of 25\%, and a system that ceases operation as soon as the toll period ends on the shorter route, that is, just before the first bulk departure (we call this the "bulk-off" system). We will use the same route 1 incident that was employed above when carrying out the analysis of the previous section. Although the downstream case did not suffer from the same problem as the upstream case we were nevertheless interested to see the impact of these modifications on costs.

We will be referring to Table 5.2, which shows summary statistics for the various combinations of toll station locations and information systems for a specific incident. The findings for this example were not representative of the general results obtained in every respect, as pointed out below. It helps to explain important features of the cost curves.
Figure 5-24: Cost (excl. tolls) - Incident on Route 1, With Info, Downstream Toll Station

Figure 5-25: Benefitogram (excl. tolls) - Incident on Route 1, Downstream Toll Station
5.3 Incidents and Information in the Toll Regime

Figure 5-26: Costs (incl. tolls) - Incident on Route 1, No Info, Downstream Toll Station

Figure 5-27: Costs (incl. tolls) - Incident on Route 1, With Info, Downstream Toll Station
Figure 5-28: Benefitogram (incl. tolls) - Incident on Route 1, Downstream Toll Station

Figure 5-29: Toll Station Location and Cost incl. tolls: No Info
5.3 Incidents and Information in the Toll Regime

![Figure 5-30: Toll Station Location and Cost incl. tolls: Info](image)

and benefitograms, however, and provides a feel for the magnitude of changes in various parameters introduced by information. In addition, 100-run simulations were used to obtain more general insight into the properties of the system.

**Upstream Toll Stations**

We begin with the case of upstream toll stations. Figure 5-31 shows the benefitogram for the incident on route 1 used in the foregoing examples, with information supplied at a 25% rate. The net contribution of the "U" is now slightly positive instead of negative as for the 100% guidance case of Figure 5-18. With 100% guidance the cost decreased from 10.18 to 10.163, as shown in Table 5.2, and 25% guidance dropped the cost including tolls to 10.1421. The number of paying commuters increased from 1003 with no information to 1028 with 100% guidance. This increase was reduced in the 25% guidance case, when only 1013 commuters were tolled. No increase was noted with the bulk-off policy as expected, in which case 1003 commuters payed.

Figure 5-32 depicts the benefitogram when information is supplied at a 100% guidance rate, but not supplied to either of the bulks. The (positive) net contribution to benefits due to the "U" was greater than for both the case of 100% and 25% guidance rates. Thus the bulk-off policy resulted in the lowest cost (including tolls) of 10.1252 (see Table 5.2).
Figure 5-31: Benefitogram (incl.tolls) - Upstream Toll Station, 25% Information

The costs without tolls remained unchanged for this example decreased from 8.5752 to 8.5308 for both 100% and 25% guidance rates. Owing to the fact that the information system was turned off just before the first bulk in the bulk-off policy gave rise to reductions that were lesser, but negligibly so, in costs excluding tolls, which increased from 8.5308 for the other two policies to 8.5312 (see Table 5.2).

We can examine the results further with the help of Figure 5-33, which compares percent improvements, in total costs including tolls, for both the bulk-off policy and the 25% guidance system with the 100% base case, and Table 5.3, which shows the range and mean values of improvements under the three systems. The first 100 runs in Figure 5-33 are with the incident on route 1, and start times of the incident increasing from left to right. The subsequent 100 runs are structured the same way but with the incident on route 2. General conclusions based on 100-run simulations include the following.

- Unlike the 100% guidance rate, negative benefits never occurred for the bulk-off policy, and only rarely for the 25% guidance system. The first observation is that both the bulk-off system and the 25% guidance are generally better than 100% guidance, with the bulk-off generating the largest benefits. The mean percent improvements are 0.615 for the bulk-off system, compared to 0.450 for the 25% and 0.204 for the 100% guidance system. The bulk-off policy never generated negative benefits, ranging from 0.035 to
Figure 5.32: Benefitogram (incl. tolls) - Upstream Toll Station, No Info for Bulks

<table>
<thead>
<tr>
<th></th>
<th>Mean w/o toll</th>
<th>Mean w/ toll</th>
<th>Paying Toll</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equilibrium</td>
<td>7.760</td>
<td>9.2826</td>
<td>1003</td>
</tr>
<tr>
<td>UPSTREAM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No info</td>
<td>8.5752</td>
<td>10.1800</td>
<td>1003</td>
</tr>
<tr>
<td>100 % info</td>
<td>8.5308</td>
<td>10.1638</td>
<td>1028</td>
</tr>
<tr>
<td>25 % info</td>
<td>8.5308</td>
<td>10.1421</td>
<td>1013</td>
</tr>
<tr>
<td>No info for bulk</td>
<td>8.5312</td>
<td>10.1252</td>
<td>1003</td>
</tr>
<tr>
<td>DOWNSTREAM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No info</td>
<td>8.5752</td>
<td>10.0013</td>
<td>881</td>
</tr>
<tr>
<td>100 % info</td>
<td>8.5308</td>
<td>9.9524</td>
<td>881</td>
</tr>
<tr>
<td>25 % info</td>
<td>8.5308</td>
<td>9.9527</td>
<td>881</td>
</tr>
<tr>
<td>No info for bulk</td>
<td>8.5312</td>
<td>9.9525</td>
<td>881</td>
</tr>
</tbody>
</table>

Table 5.2: Means for Various Combinations of Toll Station Location and Info Systems for a Particular Incident
Table 5.3: Summary Statistics for Various Systems: Upstream

1.765, whereas the 25% system benefits ranged from -0.222 to 1.70, and the 100% guidance system from -0.706 to 1.851. Figure 5-33 shows that for late incidents on route 1, both modified systems occasionally performed worse than the 100% guidance system. Late incidents on route 1 necessitate switching to route 2, which does not increase toll revenue, and may actually decrease it. The reduced switching of the modified systems is thus responsible for the lesser benefits generated. Figure 5-33 also shows that for late incidents on route 2, the modified systems performed even better relative to the base system. Again, a late incident on route 2 will produce switching to route 1, which may increase the number of toll paying commuters as the toll period on route 2 ends before that on route 1. In that case the reduced switching of the modified systems, in particular the bulk-off policy, are more effective in suppressing the rise in total cost that accompanies information in the base case.

- The total cost excluding tolls showed little variation between the three systems. The average gains were greatest for 100% guidance at 0.752%, followed by the 25% system at 0.716% and the bulk-off policy at 0.696. The ranges were also similar, namely [0.069, 2.284] for 100%, [0.069, 2.185] for 25% guidance, and [0.0572, 2.279] for the bulk-off policy. Figure 5-34, which plots the difference in improvements of the 25% and bulk-off systems compared to the 100% for 100 incidents each on route 1 and 2, shows that except for late incidents on either route, the differences were insignificant. Incidents that occur early permit the 25% system to generate similar benefits albeit over a somewhat longer period (see Figure 5-11 for the untolled case). This is not possible for later incidents, and thus the 25% and bulk-off system generate lesser benefits in that case, the bulk-off system being slightly worse because the information system is completely turned off rather than just reduced in penetration.

Downstream Toll Stations

Table 5.2 shows that, without information, the cost when toll stations are upstream climbed from the no-incident equilibrium figure of 9.2826 to 10.180, whereas it rose only to 10.0013
Figure 5.33: Bulk-off and 25% guidance compared to 100% guidance: Incl. Tolls, Upstream Toll Station
Figure 5-34: Bulk-off and 25% guidance compared to 100% guidance: Excl. Tolls, Upstream Toll Station
5.3 Incidents and Information in the Toll Regime

in the downstream case. As discussed above, this is due to the reduced flow through the tolls station in the downstream case. When the charging points are downstream, we obtain, for the 25% guidance system, a reduction in total cost including tolls from 10.0013 to 9.9527, which is negligibly higher than the 100% guidance case of 9.9524. With the bulk-off policy, costs were almost identical at 9.9525. In both cases the number of commuters passing the toll station during the toll period remained unchanged at 881, which is the number paying even without information, as compared to 1003 in the equilibrium case.

When toll costs are excluded, costs when no information is supplied were the same for both upstream and downstream locations, as expected. Improvements over the no-information case were similar to the upstream case in that both 100% and 25% guidance resulted in a cost of 8.5308, down from 8.5752. With the bulk-off policy, costs were negligibly higher at 8.5312.

Figure 5-35 compares the 25% and the bulk-off system with the 100% base system for costs including tolls, and Figure 5-36 does the same for costs excluding tolls. General conclusions based on 100 run simulations include:

- The main observation is that the modified systems are less effective than the 100% guidance system. Thus the range of improvements for the 100% system is [0.0425, 1.831], while it is [0.0425, 1.766] for the 25% system and [-0.195, 1.826] for the bulk-off system (see Table 5.4). The mean improvement is 0.641 for the 100% system, 0.611 for the 25% system, and 0.543 for the bulk-off system. Moreover, the discrepancy in the performance increases with incident start time. As in the case of upstream station locations, later incidents penalize the fact that the modified systems provide information to fewer travelers; this is particularly significant for the bulk-off policy, which gives 100% guidance until the time of the first bulk departure, and is then turned off, occasionally causing increased costs including tolls - not because of an increase in toll paying travelers, but because costs excluding tolls themselves increased. Similar to the upstream policy in that later incidents make the modified systems worse than the base system, the downstream scenario differs in the fact that for earlier incidents the modified systems perform somewhat worse than the base system. This may be because the toll period end times differ only slightly, and hence the system suffers less from the effects of the switching between the bulks. The modified systems reduced information may thus outweigh the potential benefits due to reduced switching.

- As a result of the above, the relative performance of the different systems was similar when toll costs were excluded. 100% and 25% systems were again similar, the range of benefits being [0.069, 2.284] and averaging 0.754 for the former and a range of [0.069, 2.284] with mean of 0.717 for the latter, as shown in Table 5.4. The bulk-off system was somewhat less effective, the range of benefits being [-0.189, 2.270] and the mean
### Table 5.4: Summary Statistics for Various Systems: Downstream

<table>
<thead>
<tr>
<th></th>
<th>Range &amp; Mean w\ Tolls</th>
<th>Range &amp; Mean w\o Tolls</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 % info</td>
<td>[0.0425, 1.831], 0.641</td>
<td>[0.069, 2.284], 0.754</td>
</tr>
<tr>
<td>25 % info</td>
<td>[0.0425, 1.766], 0.611</td>
<td>[0.0690, 2.185], 0.717</td>
</tr>
<tr>
<td>Bulk off</td>
<td>[-0.195, 1.826], 0.543</td>
<td>[-0.189, 2.270], 0.631</td>
</tr>
</tbody>
</table>

0.631. Figure 5-36 shows a similar worsening of the modified systems compared to the base case for later incidents, which is explained as in the case when tolls are included.

In summary, the upstream toll station location scenario generally improved as a result of modifying the information system. In contrast, the modified systems actually made things somewhat worse in the downstream case. The main reason for poor performance of the information system in the upstream case was excessive switching that caused substantial increases in toll revenue, and this was effectively reduced by both modifications. In the downstream case, less opportunities existed to switch in a manner that significantly raises the number of toll paying commuters, and the information system in its 100% guidance form did not increase total costs including tolls. The modifications to the information system thus had a generally negative effect in the downstream case, as their main drawback, namely reducing the number of travelers receiving information, was not outweighed by the curtailment of toll revenue observed in the upstream case. It appears that the observed problem with upstream locations may occur in practice for any toll station location when combined with toll periods that allow revenue to increase when switching occurs.

#### 5.3.3 Modifying the Pricing Scheme

Several modifications to the pricing scheme are now explored. First, the congestion pricing scheme is further constrained for enhanced implementability. The step-toll used in the previous chapter already reflects certain restrictions on the toll structure that are viewed as necessary for feasibility. The additional requirement that the toll structures are *reasonable* and *identical* across routes is imposed here, the optimal tolls for that case derived, and the consequences analyzed. Second, the free alternative which arises in general with facility-based pricing is explored by means of allowing tolls to be charged on one route only.

Each section begins with the motivation behind the proposed modification. This is followed by a description of the method used to derive optimal tolls under the new conditions, and concludes with a discussion of the results obtained when simulating the new patterns in the presence of incidents and information.
Figure 5.35: Bulk-off and 25% guidance compared to 100% guidance: Incl. Tolls, Downstream Toll Station
Figure 5-36: Bulk-off and 25% guidance compared to 100% guidance: Excl. Tolls, Downstream Toll Station
Equal Tolls and Reasonable Parameter Values

In this section the effect of imposing two additional sets of constraints on the step-tolls used in the previous chapter are examined. The first is the requirement that the tolls are "identical" on both routes, such that the toll period begins and ends at the same time, and that the tolls charged during the period have the same value\(^2\). The reasons for this requirement, \textit{a priori}, can be pre-scribed. For example we may wish to counter the perception that one route is treated "harsher" than the other by the responsible agencies. This relates to another potential implementability concern: Although unequal tolls arise in order to offset differences in route free-flow travel times and capacities, the differences in toll values (or toll periods) may induce shifts to the route that is perceived as superior. Throughout this research we are assuming travelers are "rational", and thus do not model this possible behavior. If it exists, however, then the analysis in this section provides insight into the effect of a measure to avoid it.

The second requirement is that the toll start and end times, and the toll charge, have \textit{reasonable} values. By this we mean that they should be easy to recognize and memorize, falling on nice round figures. For the toll period we will stipulate that the endpoints fall on a quarter of an hour (i.e. 7:30, 7:45). The toll value should be a multiple of 5 cents as the "unit" (i.e. \$2.25, \$2.30, \ldots), since this is the smallest unit used on turnpikes in the US. It appears that with electronic debiting systems the reasonability constraints are less problematic, since cash is not required, and any figure can be deducted from the card; nevertheless a whole figure seems more user-friendly, and less pretentious.

We find optimal identical and reasonable (IR) tolls for two cases. In the upstream case the tolls are imposed to be identical at point \(B\); in the downstream case identical tolls are imposed at \(C\). As will be seen, the two cases have different implications.

\textbf{Computing the Flow Patterns} In order to solve for the optimum toll we used complete enumeration over the set of reasonable tolls to find that which results in the lowest total cost. For each candidate toll \(T_R = \tau_r, t^+_r, t^-_r\), we first find the route split which results in equilibrium when this toll is applied on the two routes. This is performed by equating the equilibrium costs on the two routes as functions of the flows on each route, which can be expressed using the costs of the first departing commuter as:

\begin{align*}
\alpha T^f_1 + \beta(t^* - t^1_q - T^f_1) &= \alpha T^f_2 + \beta(t^* - t^2_q - T^f_2) \\
\alpha(T^f_1 - T^f_2) - \beta([t^1_q + T^f_1] - [t^2_q - T^f_2]) &= 0
\end{align*}

\(^2\text{As a result, there will be no uniform toll such as the one used in the previous chapter to achieve an optimal route split. Such an arrangement is comparable in some cases to a cordon pricing scheme.}\)
<table>
<thead>
<tr>
<th>Toll Station Loc.</th>
<th>N1</th>
<th>N2</th>
<th>EC</th>
<th>Toll($)</th>
<th>T⁺</th>
<th>T⁻</th>
<th>TotalCost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upstream (B)</td>
<td>1624</td>
<td>576</td>
<td>9.36</td>
<td>2.95</td>
<td>6:30</td>
<td>7:45</td>
<td>17768</td>
</tr>
<tr>
<td>Downstream (C)</td>
<td>1624</td>
<td>576</td>
<td>9.35</td>
<td>1.90</td>
<td>7:15</td>
<td>8:15</td>
<td>18499</td>
</tr>
</tbody>
</table>

Table 5.5: Optimal tolls for identical and reasonable toll structures

<table>
<thead>
<tr>
<th></th>
<th>Mean w\o toll</th>
<th>Mean w\ toll</th>
<th>Paying Toll</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UPSTREAM</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equilibrium</td>
<td>8.0869</td>
<td>9.3608</td>
<td>950</td>
</tr>
<tr>
<td>No info</td>
<td>8.8824</td>
<td>10.1562</td>
<td>950</td>
</tr>
<tr>
<td>100% info</td>
<td>8.8514</td>
<td>10.1253</td>
<td>950</td>
</tr>
<tr>
<td><strong>DOWNSTREAM</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equilibrium</td>
<td>8.4082</td>
<td>9.3375</td>
<td>1076</td>
</tr>
<tr>
<td>No info</td>
<td>9.2265</td>
<td>10.0556</td>
<td>960</td>
</tr>
<tr>
<td>100%</td>
<td>9.1972</td>
<td>10.0263</td>
<td>960</td>
</tr>
</tbody>
</table>

Table 5.6: Identical and Reasonable Toll Structures: Statistics for an Example

For any step-toll equilibrium we can express $t_q$ in terms of the toll structure $\tau_r, t^{+}_r, t^{-}_r$, using Equations B.14 and B.17 for the cases of upstream and downstream toll station location respectively, and appropriately modifying both the departure times (e.g. $t_q, \tilde{t}$), and the toll variables $t^{+}, t^{-}$ in these equations to account for the fact that the free-flow travel time is not zero, and that the upstream toll stations are not just before the bottleneck, but rather an amount of time $T_{TS}^{f}$ upstream of it. We can then solve for the equilibrium route split by substituting for $t_q$ and solving for $N_1$.

Next we determine if the toll structure is feasible, by checking whether constraints given by Equations B.21 and B.25 are satisfied. In summary, we first define the set of reasonable tolls using the restrictions on toll level and time periods, and defining ranges for these parameters. Then for each case of toll station location we perform the following: For each IR toll we (1) find the equilibrium route split, (2) check whether feasibility conditions are satisfied, and (3) compute total costs if the conditions are satisfied. The optimal reasonable toll is simply the toll with least total cost. The results of using this approach are presented in Table 5.5, and can be compared to the base case presented in Table 5.1.

**Upstream Toll Station Location** Figure 5-37 shows the cost curves with information for the incident on route 1 discussed earlier. The bulks now occur simultaneously, and considerable switching occurs between the two bulks. The net contribution of the switching between the bulks that is seen to occur at the far right of the benefitogram in Figure 5-38 is
positive, as shown in Table 5.6, where costs have decreased from 10.1562 to 10.1253. This table also shows that the number of toll paying commuters remains unchanged at 950 with information.

Imposing the requirement that the toll structures are identical at $B$ implies that no traveler can, as a result of guidance, incur toll costs that differ from the traveler’s default toll costs. This contrasts with the case discussed in the previous sections, where the information system at both 100% and 25% guidance rates caused some commuters to switch to a route of higher/lower cost, since both tolls and periods differed on the two routes. We can therefore conclude at once that providing information when toll structures are identical and toll stations are located upstream makes information provision similar to the no-toll case, since the tolls will always drop out when comparing total costs on the two routes. Benefits when tolls are included will thus equal those excluding tolls, which holds for the example, where the drop from 8.8824 to 8.8514 equals the drop in costs including tolls. Does it also follow that these benefits should always be positive? In the previous sections we found that for both the untolled and the tolled regimes, information invariably resulted in improvements in costs excluding tolls, and it seems reasonable to assume that the difference in the toll structure will not affect this “property” of information provision. To put this conjecture to the test we performed a 100-run simulation, the results of which are shown in Figure 5-39. The improvements in total cost were positive in all cases, suggesting that the problem associated with providing information to the bulk observed previously may be overcome through the imposition of identical toll structures.

**Downstream Toll Station Location** For the downstream case we only show the benefitogram for the particular incident examined (see Figure 5-40), as the cost curves are straightforward. Note that the net contribution to the benefitogram due to the bulks is essentially zero. The overall net benefits are positive in this example, as shown in Table 5.6 by the decrease in costs including tolls from 10.0556 to 10.0263. A 100-run experiment showed that the improvements in total cost including tolls were always positive, as shown in Figure 5-41. The benefits in costs excluding tolls were practically identical to those with costs including tolls. Unlike the upstream case, the departures on the two routes do not have identical start and end times relative to the toll period, despite the latter’s identical structure, because the distance from the origin to the bottleneck differs. Thus for instance the bulks depart at slightly different times, unlike the upstream case where these times are identical (and the costs excluding and including tolls are equal by definition). Nevertheless, since the differences are small, the costs are almost identical.
Figure 5-37: Costs (incl. tolls) - Info, Upstream IR Tolls

Figure 5-38: Benefitogram (incl. tolls) - Upstream IR Tolls
5.3 Incidents and Information in the Toll Regime

Figure 5-39: Improvement over No-Info (incl. tolls) for IR Tolls, Upstream

Figure 5-40: Benefitogram (incl. tolls) - Downstream IR Tolls
Figure 5-41: Improvement over No-Info (incl.tolls) for IR Tolls, Downstream

Evaluating a "Free" Alternative

One of the distinguishing features of facility-based pricing vis-a-vis zone pricing is that there are "free" alternative routes available in addition to the tolled facilities. This will probably be the case at least initially during the incremental implementation phase but could be true in general. This characteristic has not been analyzed so far in this research because our two-route network permitted tolling on both routes, and because the total demand for travel on the two routes was assumed inelastic, and thus no opportunities existed to switch to other untolled routes. In this section we attempt to relax this restriction by allowing tolls to be applied on only one route.

Computing the Flow Patterns The approach followed here to compute the flow patterns when only one route is tolled consists of two steps: First we find the route split that minimizes total travel cost, then we apply a uniform toll to the route that has a higher equilibrium cost when the optimal route split is applied. For this to work, the route with the higher equilibrium cost must be the tolled route so that the uniform toll may be charged.

We began by finding the route split which minimizes the total travel cost over both routes. This is given by the sum of (i) the equilibrium cost on the untolled route multiplied by the number of travelers on that route and (ii) the equilibrium cost on the tolled route
### Table 5.7: Optimal route splits ignoring uniform toll constraint

<table>
<thead>
<tr>
<th>Tolled Route</th>
<th>N1</th>
<th>N2</th>
<th>EC1</th>
<th>EC2</th>
<th>Toll</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1742</td>
<td>458</td>
<td>9.52</td>
<td>8.26</td>
<td>3.47</td>
<td>17506</td>
</tr>
<tr>
<td>2</td>
<td>1468</td>
<td>732</td>
<td>8.72</td>
<td>10.78</td>
<td>3.79</td>
<td>19423</td>
</tr>
</tbody>
</table>

multiplied by the number of users minus the toll revenue on that route. Using the case where route 1 is tolled for illustration, the expressions for net total costs on tolled and untolled routes imply that we need to minimize $TC$, which is given by

\[
TC = N_2[\alpha T_2^f + \delta N_2 \frac{s_2}{s_2}] + N_1[\alpha T_1^f + \delta N_1 \frac{A}{2s_1}]
\]

\[
= (N - N_1)[\alpha T_2^f + \delta(N - N_1) \frac{N_T}{s_2}] + N_1[\alpha T_1^f + \delta N_1 \frac{A}{2s_1}]
\]

expressing $TC$ as a function of the flow on the tolled route. This is a quadratic in $N_2$. After simplifying, and using the symbol $N_T$ in place of $N_1$ it yields

\[
TC = N\alpha T_2^f + \delta N_2^2 \frac{N_T}{s_2} + [\alpha T_1^f - \alpha T_2^f - \frac{2\delta N}{s_2}]N_T + [\delta(\frac{1}{s_2} + \frac{A}{2s_1})]N_T^2
\]

(5.3)

Note that the coefficient of the quadratic term is positive. This is also true when route 1 rather than route 2 is tolled. Thus total cost is a convex function in the number of travelers on the tolled route, $N_T$.

It turns out, however, that in the case of route 1 being tolled, the equilibrium cost is lower on (untolled) route 2, which means that a uniform toll would have to be applied to route 2 in order to achieve equilibrium, which is of course not possible since the route is to remain “free”. The same result held when route 1 was free and tolling was allowed on route 2. These results are summarized in Table 5.7, where $N1, N2$ are the numbers of travelers on routes 1 and 2 corresponding to the optimal route splits, respectively, and $EC1, EC2$ denote the resulting equilibrium costs. Note that in the first case, $1.26$ would have to be charged as a uniform toll on the free route for equilibrium, and $2.06$ in the second case.

These results can be explained intuitively as follows: Since the equilibrium costs on the free route are entirely made up of travel costs, their share of total costs should be minimized, and hence more travelers diverted to the toll route, where toll revenues serve to reduce the travel cost portion of the total cost. Thus, compared to the equilibrium route splits with no uniform tolls on either route, which are 1653/547 when route 1 only is tolled, and 1616/584 when only 2 is tolled, more travelers are sent to the tolled route, resulting in a split of 1468/732 when 2 is tolled and 1742/458 when 1 is tolled. This can also be seen...
Table 5.8: Equilibrium Route Splits with zero uniform toll on tolled route

<table>
<thead>
<tr>
<th>Tolled Route</th>
<th>N1</th>
<th>N2</th>
<th>EC1</th>
<th>EC2</th>
<th>Toll</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1653</td>
<td>547</td>
<td>9.18</td>
<td>9.18</td>
<td>3.29</td>
<td>17711</td>
</tr>
<tr>
<td>2</td>
<td>1616</td>
<td>584</td>
<td>9.18</td>
<td>9.18</td>
<td>3.02</td>
<td>19675</td>
</tr>
<tr>
<td>None</td>
<td>1634</td>
<td>566</td>
<td>9.38</td>
<td>9.38</td>
<td>0.00</td>
<td>20638</td>
</tr>
</tbody>
</table>

by substituting the route splits corresponding to either both routes tolled or untolled into

$$\frac{\partial TC}{\partial N_T} = N\alpha(T_U^I - T_T^I) - \frac{2\delta N}{s_U} + 2N_T[\delta\left(\frac{1}{s_U} + \frac{A}{2s_T}\right)]$$  (5.4)

where subscripts $U$ and $T$ refer to tolled and untolled respectively, which yields a derivative of total cost with respect to commuters on the tolled route of -3.41 when route 1 is tolled, and -2.37 when only 2 is tolled. This property of the optimal solution implies that the route costs are higher on the tolled and lower on the untolled route vis-a-vis the appropriate equilibrium solution\(^3\). The total cost function is depicted for the case when route 1 only is tolled in Figure 5.42.

Since the untolled route will have lower equilibrium costs corresponding to the optimal route split than the tolled one, a uniform toll would have to be applied to the untolled route to attain equilibrium. This is not allowed, since the uniform toll can be positive only on the tolled route. As the total cost function is ditonic, it seems to follow that at the optimal solution the uniform toll on the tolled route is zero. Table 5.8 shows the results for this case, where we simply find the route split at which the two routes are in equilibrium without the influence of uniform tolls. Charging either route results in improvements over the situation when both routes are untolled, which is also listed in the table.

Recall that the total cost function is based on the optimal step-toll, which minimizes travel costs excluding tolls on a route for a given route split. This is an appropriate procedure only if we can achieve any route split. This was true for the unrestricted tolls on both

\(^3\)The findings for our example might not always hold. In general, when no uniform toll is applied to either route the equilibrium number of travelers on the tolled route, $N_T$, is given by

$$N_T = \frac{N\frac{A}{s_U} + \alpha(T_U - T_T)}{\delta\left(\frac{1}{s_U} + \frac{A}{2s_T}\right)}$$  (5.5)

which might not always be less than the optimal number of travelers of the tolled route, $N_T^*$, which is given by

$$N_T^* = \frac{N[\frac{2A}{s_U} - \alpha(T_U - T_T)]}{\delta\left(\frac{1}{s_U} + \frac{A}{s_T}\right)}$$  (5.6)
5.3 Incidents and Information in the Toll Regime

Figure 5-42: Total Costs as a function of flow on tolled route 1

routes examined previously, where a uniform toll could be applied to either route, without affecting total cost, to obtain the optimal route split. It is therefore conceivable that in this case a better solution exists if the toll is allowed to be sub-optimal (with regard to total cost on the tolled route for a given flow on the route), since a "better" (in the sense of overall contribution to total cost on both routes, in particular greater than the equilibrium splits just presented) route split may be achieved. Great improvements over the solution just presented are not to be expected, however, since the primary benefits from pricing come from departure time, not route, adjustments.

Thus to solve for the true optimal toll for this problem, it is required that the uniform toll is zero on the free route and on the tolled route, but the toll may be sub-optimal. We have performed a fine-grained search over all three toll variables, namely the start $T^+$ and end $T^-$ of the toll period, and the toll level, $\tau$, using intervals of one minute for the time endpoints and one cent for the toll level, and using broad ranges for the variables ($1.00 - \$5.00$ for the toll level, and time-brackets wide enough to capture all possible feasible tolls). An even finer search was carried out in the vicinity of the toll that resulted in the lowest value thus obtained. Due to the complexity we have limited ourselves to examining the case where route 1 is tolled, which is by far the more fruitful choice in terms of expected benefits. This proved computationally very intensive, demanding several hours of CPU time on a Cray X-MP EA/464.
Chapter 5  Combined Pricing and Information Systems: Analysis

It turns out that the truly optimal toll for this example coincides with the simple approach carried out above, namely that the route split corresponds to zero uniform tolls, and the step-toll is optimal in departure times. It was also observed that it is easy to find solutions that differ significantly in one or more of the toll variables, yet are reasonably close to the optimal solution in terms of total cost. This suggests considerable flexibility in designing practical tolls that are both efficient and satisfy specific objectives of the toll authority, such as revenue targets or limits on the length of the toll period.

Results and Discussion  Examples of simulation results for the optimal toll are shown in Figures (5-43 to 5-45) for the case of a route 1 incident. Figure 5-43 is straightforward - the costs on route 2 reflect a no-toll equilibrium without incidents, such as we have seen at the beginning of the chapter. The cost curve for route 1 is also familiar, showing an incident acting on a toll-equilibrium. Note that the last departure on route 2 occurs after the last departure on 1.

Figure 5-44, which shows the with information case, is similarly straightforward. The benefitogram in Figure 5-45 shows that the total benefits are quite high, even with the dip in benefits arising from the batch of commuters whose default route is 2 and who leave after the end of the toll period. But when the incident is on the untolled route, the benefitogram is correspondingly negative to a large degree, as shown in Figure 5-46. In contrast, when both routes were tolled the negative contribution to the benefitogram due to the bulk on the incident route was more or less balanced by the positive contribution of the bulk on the incident-free route. In the case of one tolled route, benefits can only balance out over a number of days when incident occurrence on both routes. The results of testing 100 incidents on each of route 1 and 2 are displayed in Figures 5-47 and 5-48 for costs including tolls. Benefits are always significant when the incident is on route 1, as was the case for the example. When the incident is on route 2, most runs yielded negative benefits, except for incidents that were very early or late. In those cases the toll period has either not started, or has ended, as switching from route 2 to route 1 occurs. Thus the problematic increase in the number of toll paying commuters, which occurs at other times, is avoided.

Since negative benefits occur only on route 2, one might contemplate improving this situation by making the information system activation sensitive to the location of the incident. In that case whenever the incident is on route 2, no information is provided. This would avoid the negative benefits associated with that case. It would probably be considered unfair, however, by the travelers on route 2, who benefit overall from information when the incident is on route 2 even though net benefits for travelers on both routes are negative.
5.3 Incidents and Information in the Toll Regime

Figure 5-43: Costs (incl. tolls) - No Info, Toll on Route 1, Incident on 1

Figure 5-44: Costs (incl. tolls) - Info, Toll on Route 1, Incident on 1
Figure 5-45: Benefitogram (incl. tolls) - Toll on Route 1, Incident on 1

Figure 5-46: Benefitogram (incl. tolls) - Upstream Toll on Route 1 only, Incident on 2
5.3 Incidents and Information in the Toll Regime

Figure 5-47: Improvements in Costs (incl. tolls) Incident on 1

Figure 5-48: Improvements in Costs (incl. tolls) Incident on 2
5.4 Testing the Superadditivity Hypothesis

In the previous sections we have shown the impact of information on both tolled and untolled regimes. For the case of toll regimes, the information system was "optimized" to result in improved performance and avoid the phenomenon of negative benefits that occasionally occurred when the route guidance system operated in its standard mode. Almost all of the scarce literature on joint pricing and information (i) contemplated the possibility of synergistic effects and (ii) suggested that they would result because of the way the two systems were to be implemented, namely through making pricing effectively a means of engendering compliance to the guidance produced by the information system. As discussed in the previous chapters, we think this is inappropriate from an implementability point of view. Nevertheless, the question of synergism remains of interest when asked in the context of the implementable systems to which we are restricting our attention.

What can be said a priori about whether gains are superadditive or not in our case? Our informal argument in Chapter 1 suggested that recurring congestion would be lower in the presence of congestion pricing and there would be more excess capacity on routes, making rerouting more effective, suggesting that the policies are superadditive.

First, note that with the model of pure bottleneck congestion that we are using in this research, capacity is not a meaningful concept. Traditional models of congestion, such as the BPR curves (see Sheffi [1985]), which employ the notion of capacity are also not very persuasive, since capacity is defined as flow per unit of time such as an hour or 15 minutes, an average concept that is ill-defined in real-time. However, if we assume, along the same lines, that as a result of congestion pricing, queue lengths on average are shorter than with the unpriced regime, then, based on Appendix D, it is more likely that the number of trapped commuters on link 1 or 2 is less with pricing than without. Hence, benefits to both the initial and subsequent packets are reduced. This suggests that information is less beneficial with pricing, i.e. the policies are subadditive.

We examine the superadditivity hypothesis based on total cost excluding tolls, because the objective of congestion pricing is to minimize travel costs, and tolls are considered a transfer payment. Total travel costs excluding tolls were found to be rather insensitive to the variations in the information regimes tested, and hence we evaluate the superadditivity hypothesis for combined pricing and the base information system. Figure 5-49 shows the results of a 100-run simulation when the incident is on route 1, with the horizontal axis representing the runs in increasing order of start time of incident. Note that all points are negative, indicating subadditivity. Figure 5-50 shows the same plot in addition to the benefits due to pricing, information, and pricing and information out of which it is composed. Note that the magnitude of the subadditivity points is small compared to the other points, suggesting weak subadditivity. In other words the difference between the
gains due to information for the untolled regime and the gains due to information in the tolled regime are small compared to the gains themselves. With the incident on route 2, we obtain the plot of Figure 5-51, in which all points but one are positive, indicating superadditivity. Once again the magnitude of the superadditive points is small compared to the gains due to information itself, as shown in Figure 5-52. Why is there subadditivity if the incident is on route 1 and superadditivity when it is on route 2? One possible explanation is as follows. The route splits in the tolled regime (1620 and 580 commuters on routes 1,2 respectively) are such that there are fewer travelers on route 1 and more on 2 compared to the untolled regime (1634, 566). Thus there are more travelers on route 2 in the tolled regime. As a result, information may be able to generate greater benefits when the incident is on that route compared to the untolled regime. To investigate this further, we examined the superadditivity hypothesis when the route split in the toll regime was held at (1634, 566), the value for the no-toll regime. The results are shown in Figures 5-53 and 5-54 for incidents on routes 1 and 2 respectively. When the incidents is on route 1, points of both superadditivity and subadditivity exist, and their absolute values are very small. With the incident on route 2, the maximum absolute values are somewhat larger but are also distributed both above and below zero. These results appear to support the explanation offered above. It thus seems that in the presence of pricing, the information system will generate benefits that are practically the same as those achieved in the no-toll regime. This may be due to the linear nature of the cost functions, which is not affected by the starting point (i.e. the regime).

5.5 Summary

This chapter has examined the combined implementation of congestion pricing and traveler information systems using a combination of analytical models and simulation. The discussion began with an examination of the case in which information was provided to the untolled regime. Results were found to be in general agreement with those of previous work in this area. In addition we were able to conclude that benefits increased with guidance rates for the information system used here, which is assumed to possess perfect prediction capabilities.

In regard to evaluating the joint implementation of congestion pricing and driver information systems, we first investigated the impact of toll station location. In the absence of information the costs excluding tolls for the upstream and downstream locations are identical since route choices are unchanged.

Negative benefits due to information in total costs including tolls were occasionally observed in the upstream scenario. This occurred primarily when information provision resulted in a net increase of toll-paying commuters. It should be recalled that in the priced
Figure 5-49: Evaluating Superadditivity: Incident on route 1

Figure 5-50: Evaluating Superadditivity: Incident on route 1, Component Curves
5.5 Summary

Figure 5-51: Evaluating Superadditivity: Incident on route 2

Figure 5-52: Evaluating Superadditivity: Incident on route 2, Component Curves
Figure 5-53: Evaluating Superadditivity: Incident on route 1, No-Toll Route Splits

Figure 5-54: Evaluating Superadditivity: Incident on route 2, No-Toll Route Splits
5.5 Summary

regime, guidance is always based on total costs including tolls.

In the downstream case, and without information, total costs including tolls were frequently less than for the upstream case. This occurred whenever the incident overlapped with the toll period on the respective route, since the slower flow through the toll station reduces the number of toll-paying commuters. When information was provided, positive benefits were always observed for total costs excluding and including tolls. Unlike the upstream case, the net change in the number of toll-paying commuters was negligible in the downstream scenario, and this appears to overcome the problem of negative benefits. Another feature of downstream toll station locations is that whenever an incident begins before the beginning of the toll period, commuters may be drawn into the toll period who do not usually pay the toll in the no-incident case. In summary, the total cost including tolls was generally lower than for the upstream case, both with and without information. The use of the cost curves and benefitograms, which exploit the deterministic structure of our model by keeping track of costs for individual travelers under various scenarios, helped to detect the finer points just made. Looking at only the aggregate fact that information produced benefits might have obscured some of the effects observed.

Two modified information systems, one with a 25% guidance rate and one in which the information system is turned off just before the first bulk, were evaluated as well. The problem of negative benefits due to information was solved most effectively by the bulk-off policy.

This chapter also considered imposing restrictions on the toll structures that require that they are identical across alternative routes and that the toll variables (start and end times and tolls charged) are reasonable and easily recognizable. We derived these tolls and examined their impact when information is supplied. For the upstream case, the problem of increased costs due to providing information was overcome with the new toll structure.

Another modification to the pricing system which we discussed is the case where only one road can be tolled. It turned out that finding the optimal toll was not quite as straightforward as might be expected, since ideally one would like more commuters to use the tolled route than is possible with the restriction on keeping one route untolled. Simulation results when the tolls were in place indicate that benefits may be negative when the incident is on the untolled route. Possible steps to take account of this may include that of making the information system activation sensitive to the location of the incident, which may however not be acceptable to the public.

Finally, we evaluated the superadditivity hypothesis that is frequently discussed in the literature. It was found that total benefits due to pricing and information were slightly subadditive for incidents on route 1 and slightly superadditive for incidents on route 2. This was explained as a consequence of the route splits in the toll regime, which are different compared with the untolled regime. Thus the higher number of commuters on route 2 in
the tolled regime resulted in slightly higher benefits due to information compared to the untolled case, and a similar argument explains the subadditivity for route 1 incidents. We tested for superadditivity when the route splits in the toll regime were identical to those in the untolled case, and observed essentially simple additivity, supporting our interpretation.
Chapter 6

Conclusions and Future Research

6.1 Summary

This thesis has three parts. In part 1 we introduced the topic of this thesis, discussed the joint implementation of congestion pricing and traveler information systems, reviewed the literature, and characterized both pricing and information schemes. This part culminated in a discussion of various design features with a view to making specific recommendations as to what approach should be followed in practice (considering both acceptability and operability). The issue of public acceptability of congestion pricing was discussed, and the concepts of fairness theory used to compare the acceptability of pricing with that of other traffic restraint measures. (Chapters 1, 2, and 3 and Appendix A)

In part 2, we identified various design issues that apply to integrated pricing and information systems. The decision to choose a simple model to address some of these issues was explained, and the model described. The model chosen utilizes existing equilibrium results to derive default travel patterns, and uses simulation to incorporate the occurrence of incidents and the traveler information system. The equilibrium model was re-derived and an error in one of its results corrected. (Chapter 4 and Appendices B and C)

Part 3 was devoted to quantitative analysis. We began by describing the impact of discretization (using simulation) on an analytical (closed-form) model. The impact of information in the presence of incidence on the untolled regime was then examined. Results were found to be consistent with those of earlier research. We described cost curves and benefitograms, both of which were used extensively to analyze the behavior of travel costs under pricing and information. This lead to an investigation of the central theme of this thesis, the interaction of pricing and information systems. First the impact of toll station location was discussed. The consequences of providing information to the departure "bulks" that occur at the end of the toll periods was also studied, and it was found that total costs including tolls can increase as a result of information if toll stations are located upstream.
In that case, negative benefits were observed when switching due to information raised total toll revenue. Just how information generates benefits despite the lack of added capacity is analyzed in Appendix D.

Modifications to the information system were then tested. The upstream scenario generally improved with the modified systems, overcoming the problems of negative benefits. The downstream case showed some deterioration.

Modifying the pricing system was explored next. We began by deriving tolls that have identical structures across routes, and which have reasonable values that are easily recognized by travelers. The identical toll period end times helped to alleviate the problem of negative benefits in total costs including tolls observed previously. Another modification was to allow only one route to be tolled. We derived the optimal tolls both using an approximate and an exact approach. It was observed that the net benefits were highly positive/negative depending on whether the incident occurred on the tolled/untolled route (Chapter 5).

Finally, superadditivity was examined, and it was concluded that for our model, only simple additivity (independence) appears to hold.

6.2 Conclusions

The main objective of this thesis was to evaluate the joint implementation of congestion pricing and driver information systems. To that end, we selected a specific pricing and information system, and developed a modeling framework to study "integrated" design issues that pertain to the interaction between the selected systems. The main findings of the analysis of these issues are summarized in the previous section. Despite being an advance over earlier work, the models used are still very simple, making several strong assumptions. The extent to which the conclusions can be generalized and bear on actual policy is now discussed. We focus on three issues: Toll station location, negative benefits with information, and superadditivity.

Several phenomena were observed to differ between upstream and downstream toll station locations, including that of benefits in total costs including tolls. We have attempted to capture the impact of the natural travel time variability by comparing an upstream location (relative to the points of possible congestion) to a downstream one. In our model the upstream location exhibited zero variability in arrival times at the toll stations. For the downstream locations, incidents may introduce uncertainty in arrival times at the toll stations, and as a result, uncertainty in toll costs. The main findings that directly relate to that distinction is that (1) downstream toll station locations may distort toll costs by delaying travelers "into" the toll period and pushing some out (compared to their default toll costs), and (2) the reduced flow that occurs downstream of an incident may reduce
the total number of toll paying travelers. These findings appear to hold in general, since in practice incidents do entail reduced flows as a result of reduced capacity, which is also borne out by congestion functions that differ from the deterministic bottleneck used here.

This potentially poor predictability of travel costs can be reduced by the information system. However, even with an information system, a portion of travelers may not be able to respond to information, namely those "caught between" the points of route adjustment and the incident. If predictability is seen as an important implementability issue, then toll stations for facility-based pricing should probably be located at the entrance to the facility (which means that the toll charged is based on the time of entrance). In practice the variability between a traveler's origin and the entrance to the facility may still potentially give rise to the phenomena observed with downstream locations. However this is less likely if the travelers have either only a short leg of the trip between the entrance and the facility, or if that leg suffers relatively little congestion and hence the variability in travel times (which is the actual cause) is low. Moreover even if the arrival time at the entrance to the facility is highly variable, then the reason for arriving at the "wrong time" at the toll station is at least not directly associated with the tolled facility. Again, if "reliable" (good predictability of costs) toll station locations are not possible\(^1\) but viewed as desirable, an information system may help to reduced uncertainty at intermediate nodes, as was the case in our example. This may be even more important if travelers regard tolls in a somewhat "irrational" manner that overstates their cost as compared to the other less tangible costs of travel time, and are likely to prefer less tolls even at potentially higher total cost.

The fact that total revenues may decrease in the downstream case must be taken into account by the toll authority, in particular if conflicting objectives exist with regard to the management of the toll revenues\(^2\).

Negative benefits in total costs including tolls were sometimes observed for the upstream case. Closer inspection revealed that this was linked to a rise in the total number of commuters as a result of the switching due to guidance. Therefore, it seems that this could be observed for downstream locations as well, if toll periods are such that revenue-increasing switching can occur. Beyond the confines of our model it is not clear how well this observation can be generalized. In practice, the lack of travel homogeneity might force guidance provision that excludes the use of the travelers preferences. For example, travel time and tolls could be predicted separately, and the decision of combining these costs left to the traveler. This would further complicate predicting the response to guidance. It is

---
\(^{1}\)This may be the case for some cordon pricing schemes, where the location of stations is often governed by other factors, such as the geography of the tolled area.

\(^{2}\)That is, redistribution may face competition from a desire to maximize toll revenues, possibly for other purposes than transferring the funds back to the public. See Goodwin (1930), Evans (1992), and Bernstein and Muller (1994) and for further discussion of this point.
here that the findings regarding identical and reasonable tolls become useful. With such
toll structures, which appear particularly likely to offer themselves as a reasonable choice in
the cases of “parallel” corridors, guidance can be based on (preference-independent) travel
times only, since toll costs drop out. Since these structures also preclude changes in total
revenue as a result of information, it appears possible that the phenomenon of negative
benefits that was eliminated in the model would be avoided in practice as well.

The conclusion that the policies of pricing and information are independent seems to be
a consequence of the particular congestion model assumed, namely a linear deterministic
queue. It is not clear whether it will hold with other models, and how it generalizes to the
real-world.

We feel that the findings of this research have validated the basic point of this thesis,
that there are issues specific to the combined implementation of pricing and information
systems that deserve special attention. The variety of interesting issues that were observed
and discussed should provide a powerful stimulus for further research in this area. In the
next and final section of this dissertation, some pointers for future research are outlined.

6.3 Future Research

The work presented in this thesis should be extended along several directions for increased
modeling realism. Issues that should be explored include:

- Heterogenity of commuters. We have assumed throughout that commuters have the
same characteristics, such as the value of travel time, the value of arriving early, etc.
This is clearly not true in practice. The removal of this restriction is very important
because traveler characteristics determine the response to the pricing scheme. Moreover
if information is to be provided in the form of guidance, one should fully allow
for heterogeneity, or else the toll costs on the different routes can only accompany the
predicted travel times, as discussed previously. Travelers must then combine these
costs depending on their own preferences, which in turn may make the estimation of
their response more difficult.

It appears unrealistic to expect to be able to estimate or measure the characteristics of
an entire population. Even much cruder approximation to heterogeneity have entailed
considerable modeling complexity. The work of Arnott et al. (1988), Muller (1993),
and Muller and Bernstein (1993) assumed that travelers belong to one of two groups.
Obtaining equilibria for such simplified scenarios under these conditions was found
to be very difficult. More work is therefore also needed on the effect that simplifying
assumptions about population characteristics have both on the accuracy of the
predicted response and on the impact on individual travelers.
6.3 Future Research

- **Analysis of general networks.** The simple network representation here was limited both in comparison to real-world networks, and in terms of the adjustment margins permitted, which only included route and departure time choice. As discussed above, at present no solutions exist for general networks even when limiting attention to these two choices, and ignoring other adjustment possibilities such as suppression of trips, switching to another mode, etc. Similarly, stochasticity in various forms should be introduced into the model. User perceptions of travel costs, network capacity, and the demand patterns, such as the route and departure time choices of the commuter, can all be assumed to be random variables.

- **Pricing of other externalities.** It seems reasonable to consider the pricing of congestion as only the first step in implementing comprehensive marginal cost pricing in the highway travel sector. Pollution and road damage are other costs that are incorrectly priced. Other modes of transport, primarily transit, should also be examined with regard to their pricing practices. Changes in pricing policy, however, are always likely to be politically controversial, and work on the political consequences of a particular change is itself important in that regard. Work on the development of models that are able to recommend the appropriate pricing policies for these instances also needs to be intensified.

- **Disequilibrium behavior.** The extensions discussed above are easy to identify, but represent formidable challenges to those seeking to carry them out. The extreme complexity of solving for dynamic equilibrium on general networks even with simplifying assumptions also raises the question of whether the notion of an equilibrium is a realistic one on which to base a great deal of effort. It is not clear that travelers can be assumed even merely to *tend* towards such an equilibrium when the underlying computational requirements on them may be excessive, especially in the presence of inevitable randomness. Thus more work on both travel behavior is needed. In this context the efforts of information system designers, who must learn about actual travel decisions made in response to a variety of conditions, should prove useful.

Even if equilibrium is assumed to be a realistic view of actual behavior, the theory discussed so far does not specify how long it would take to move from the unpriced to the toll equilibrium, nor what trajectory the daily total system costs will describe. Specific assumptions about travelers’ day-to-day adjustment behaviors will have to be made in order to learn more about these points. This becomes more difficult if stochastic variations, or the presence of an information or traffic management system is to be modeled as well. Disequilibrium models of general application in economics are discussed by Fisher (1983); an example of an application to idealized information.
systems is given by Friesz et al. (1994).

- **Modeling other Advanced Technologies.** This research was limited to examining the interaction between pricing and an idealized information system. It thus tended towards a best-case analysis. The interaction of more realistic information systems, or other traffic management systems, with pricing systems should be modeled.

- **Highway mode pricing.** In some cities, traffic conditions are highly congested for prolonged periods (Guiliano [1992]). Thus well-defined peak-periods do not exist, or extend through most of the day. In contrast, our model assumed a narrow peak-period, with no travelers outside of these periods to be worried about. In that case the toll period would have to be very long, with departure time adjustments to pricing becoming increasingly unattractive as they would be made at very unusual hours. The main shift would then be to other modes, i.e. the entire highway mode is priced. This approach will probably viewed with even more hostility by the automobile “lobby”.

- **Practical experiments.** Finally, the most important work to be carried out in our view in regard to congestion pricing and its combination with other technologies is to experiment through pilot projects. The inability to precisely predict the consequences of any policy change should not hamper efforts towards implementation when the general direction of changes is reasonably well established. Singapore has seen dramatic decreases in car travel during peak hours as a result of the pricing scheme established there. It is unlikely that we shall have highly accurate models at our disposal any time soon. The best way to gauge responses, including the public’s perception of the policy, and improve existing models is through analyzing data obtained from actual tests. The current work in progress in California appears to be the right step to take at this point in time.
Bibliography


[102] The Urban Transportation Monitor. Grant awarded for first congestion pricing project, October 1, 1993.


Appendix A

Public Acceptability and Perceived Equity

After reviewing some work on public acceptability, we turn to a brief discussion of fairness theory. As discussed above, many people have discussed the equity aspects of congestion pricing and have claimed that it has been a major obstacle to implementation. However the same criticism can be raised against many policies which have been implemented in the past. The purpose of this appendix is to consider what is unique about congestion pricing. We use fairness theory in an informal way to compare pricing with other traffic restraint measures. Some other aspects of equity are discussed at the end of the section.

A.1 Public Acceptability

The paucity of actual implementations has spawned a vast, and inevitably somewhat repetitive, literature that seeks to explain the consistent public opposition to congestion pricing. Recent technological developments have weakened objections based on the difficulty, and high cost, of implementing a scheme (Else [1986], Seila and Wilson [1991]), and those based on the privacy issue (e.g. Borins [1986], who analyses the Hong Kong experience). The almost exclusive attention given to congestion pricing schemes that are time-variant has made the large literature based on static models less useful, since it does not explicitly handle behavioral adjustments to the toll when departure time adjustments are possible. The literature on acceptability has become increasingly pessimistic over the years, as more and more proposed schemes were rejected even before the implementation of pilot projects. At the same time, more effort has been expended in identifying specific ways of easing the transition into pricing that would alleviate some of the problems associated with untransformed pricing schemes a la Singapore. An example is the paper by Starkie (1986), who suggests the use of two-part auctions or the introduction of a variety of money-time options.
by constructing competing tolled routes. Alternatively, Bernstein Guiliano (1992) provides a more recent discussion of acceptability. Her main conclusion is that it is unlikely that congestion pricing will be implemented to any significant extent in the U.S. because of public skepticism, resistance to high tolls, and pressures to divert tolls to new transportation facilities.

Stated preferences approaches to evaluating the potential response to pricing proposals have been performed by several researchers. Jones and Hervik (1991) report that acceptance of road pricing schemes jumped from 30% to 63% when explicit proposals were made for the use of the generated revenues.

In addition, those who argue, using theoretical models, that congestion pricing does not lead to Pareto improvements\(^1\) must remember that there is a whole family of models in which commuting costs may not only remain unchanged, but could even decline, *even if the toll costs are included*. This is the general group of models that takes into account the possibility of adjusting departure times, which are based on the work of Vickrey (1969).

### A.2 Fairness Theory

Fairness theory is an approach to the analysis of equity whose origins lie in the work of Foley (1967), and which has been extended by Pazner and Schmeidler (1974) and Varian (1974), among others. The discussion in this section is based primarily on Baumol (1986), who, in cooperation with Fischer, examined the equity of the classic peak pricing solution for cases with and without congestion. Their formal analysis deals with equity issues from the perspective of Pareto improvements. As mentioned previously, the classic peak pricing model is unrealistic in several respects when applied in the transportation context, even when congestion is taken into account. The value of their contribution, however, lies in their informal discussion of the equity of peak-period pricing in transportation from the perspective of fairness theory. The key observation is the recognition that even when peak-period pricing produces a Pareto improvement, many people may nevertheless consider it grossly inequitable. This is because the arrangement may be regarded as elitist by those who do not pay the toll but change their behavior, and arouse feelings of envy and indignation. This is the only contribution to address equity issues in congestion pricing from the perspective of formal fairness theory, which is described next.

Fairness theory differs from the analysis of Pareto improvements, in that it considers the envy between the affected parties. Fairness theory has been altogether neglected when

\(^1\)In practice, the lack of Pareto improvements is not a reason for deciding a project is not to be implemented, but rather the scale of the inequity is what is most important. In the words of Donald McAllister (1980) "there probably has never been a public action that did not make someone worse off, and there probably never will be".
discussing congestion pricing, except for the Baumol (1986) contribution.

A.3 Basic Concepts

Some of the basic concepts are described below:

A distribution of $n$ commodities is said to involve envy by individual 2 of the share obtained by individual 1 if 2 would rather have the bundle of commodities received by 1 under this distribution than the bundle the distribution assigns to 2.

Note that this definition does not rely on interpersonal comparisons of utilities. It means that if individuals $A$ and $B$ receive commodity bundles $a$ and $b$ respectively and $U^a$ and $U^b$ are their respective utility functions, envy of $B$ by $A$ is not defined as $U^b(b) > U^a(a)$, but as $U^a(b) > U^a(a)$. $A$ consults only his own preferences and decides whether he would rather have $B$'s bundle or his own, and need not consider $B$'s preferences.

A distribution is fair if it involves no envy by any individual of any other.

A distribution is strictly superfair if each participant receives a bundle that is strictly preferred by that individual to the bundle received by anyone else, that is, if his holdings could be reduced (in the case of divisibility) without giving rise to envy.

It follows from the above that a Pareto improvement need not be fair - take the case of a change that makes both $A$ and $B$ better off, but $B$ much more so, which could evoke envy in $A$.

A.4 Applying the Theory Formally

We have already reviewed the informal discussion of the fairness of peak-period pricing in transportation by Baumol (1986). Can we apply formal fairness theory in our context? The answer seems to be no, as long as we are dealing with equilibrium models of homogeneous commuters - a policy induces a change, and as a result of the change, a new equilibrium is reached, with everyone having the same new cost. The cost composition can differ among commuters, but of course inherent in such models is the assumption that it is total cost that matters, and hence commuters are indifferent to various combinations that yield the same total cost. In addition, it is assumed that there is no cost associated with the behavioral change that some commuters will in practice undergo. Thus all we can conclude from a welfare point of view is whether a Pareto improvement has occurred, which applies if the new equilibrium cost is lower than before. If that is not the case, and if we assume redistribution to occur to some degree, we will only be able to make Pareto type evaluations if the redistribution is specified in detail. Otherwise, only the broader benefit-cost criteria

---

These represent Baumol's choice and differ somewhat from the terms used by earlier authors.
A.5 Congestion Pricing vs Non-Fiscal Traffic Restraint Measures

In the same context, let us take a brief look at some recently implemented traffic restraint measures which do not employ pricing. In Milan, Italy, private vehicles are prohibited from entering a cordon surrounding the city center between 7:30 a.m. and 4:30 p.m. Similar restrictions on private cars have been imposed in Florence, Rome, and Bologna in Italy; Strasbourg, France; Goteborg, Sweden; and Tunis, Tunisia (May [1986], Jones and Harvik [1992]). Now, as Hau (1992) notes, such measures are tantamount to setting prohibitively high prices on vehicle usage whereas with pricing motorists still retain the choice of when and where to travel, at least those who can afford to do so. Thus from the same viewpoint as that used in evaluating pricing, we may conclude that these approaches are actually worse than pricing, since the entire population of previous car commuters needs to change their behavior. Why are they implemented as much, i.e. why are they politically acceptable, when congestion pricing has yet to see the light of day in a reasonably democratic system?

We suggest that the main reason is that these systems are fair in the above defined sense\(^3\). The restraint measure that bans vehicles altogether during certain times does not distinguish on the basis of income; although it may be argued that, objectively, some travelers may be hurt more than others, both in a relative and an absolute sense, it appears that the perception is that everyone suffers equally. Then the question becomes: If they all suffer, who cares about fairness - shouldn't the project meet with solid opposition? Again, we think that there is enough consciousness in many industrial states with regard to the most visible problems of congestion, such as environmental pollution and stress, that a basic willingness to do something about it exists, even if it includes some personal sacrifice\(^4\). However, history has shown, and the nature of congestion, which is a mass phenomenon, suggests\(^5\), that without some externally imposed restraint or encouragement, even well-disposed individuals are unlikely to take such steps on their own in large numbers. As for congestion pricing, it may be seen as a policy in which the only ones who sacrifice are those who actually change their travel pattern, and may thus fail to satisfy the desire to be part

---

\(^3\) Other possible reasons include the more compact structure of European cities compared to those in the U.S.

\(^4\) This point applies less to some developing countries if (i) they have less pollution or congestion to begin with (ii) they have more “urgent” social problems to have developed this consciousness, (iii) political systems tend to be less democratic i.e. policy makers tend to take on more the role of who decides what is good and equitable.

\(^5\) If I stop driving and take the bus, will it make any difference? It might even be the case that “my space” will be taken by somebody else.
of a collective effort to do good.

A.6 Further Aspects of Equity

It is sometimes said that the introduction of congestion pricing is no different than the market pricing in place in other sectors of the economy. Poole (1992), for example, remarks, with regard to the fairness (in its informal sense) of congestion pricing, that Americans are used to selecting among combinations of price and service, such as for air travel, restaurants, express mail, hotels etc., this being an ordinary, everyday phenomenon. The difference is that in such situations, the new and more expensive product is offered with generally no impact on the quality and availability of the cheaper product. In the case of congestion pricing, the better "product" (faster travel times during the "peak-period") is achieved through the behavioral adjustment of the less well-off, who incidentally may receive a product that is also better in some sense (traveling at less convenient times possibly offset by reduced travel times), but may nevertheless perceive themselves as being worse off.

Giuliano (1992) states that although equity is frequently identified as a major barrier to implementing congestion pricing, it can be argued that the distributional aspects of congestion are not particularly important with respect to public acceptance because (i) equity appears not to be a major public policy issue as federal and state tax and revenue policy has become less equitable due, among others, to the reduction and elimination of redistributive programs and because (ii) distributional considerations have not been a definitive consideration in other areas of transportation policy, such as public transit, where support exists for the notion that both transit fare and investment policy are frequently regressive. She also suggests that distributional equity may present an apparently legitimate basis for opposition that is actually motivated by other reasons. True as these points may be, this argument misses the distinction between perceived and actual (or rational) equity which we made above, namely that the main criterion for public acceptance is that of perceived equity, which may differ strongly from one individual to another even when the economic analysis suggest similarities, and which is affected by less easily measurable factors such as envy.
Appendix B

Deriving the Optimal Toll

This analysis derives the optimal toll for both the case where the toll stations are located after the bottleneck, and when they are located before it. Although the results for that case can be obtained directly from Arnott et al. (1990a), we have chosen to present here the actual derivation, which is not given in that reference and which is interesting in its own right. Furthermore, the intermediate results are necessary for carrying out the analysis of Chapter 5.

B.1 Step-Toll Equilibrium

The constant value of the toll is $\tau$. For equilibrium and under the assumption that the queue length is never zero for a positive amount of time between $t_q$ and $t_q'$, there is a departure free gap for a period of $\tau/\alpha$ which begins at that point in time that would result in arrival at the toll station location just as the toll period started. We are assuming, for simplicity in deriving the results, that the free-flow travel time anywhere between the origin and destination is zero. Note that the results in chapter 4, which do not make this assumption, are appropriately offset to take this into account. The equilibrium departure rate is the same as before, since the toll is constant within the toll period. $t^+$ and $t^-$ are the beginning and end of the toll period when the toll station is downstream of the bottleneck. $T^+$ and $T^-$ are defined as the departure times which result in an exit time from the bottleneck of $t^+$ and $t^-$ respectively, and hence are the beginning and end of the toll period when toll station location is upstream of the bottleneck. These quantities are related as follows:

$$t^+ = T^+ + \frac{D(T^+)}{s} \quad (B.1)$$

which reduces to

$$t^+ = \frac{\alpha}{\alpha - \beta} T^+ - \frac{\beta}{\alpha - \beta} t_q \quad (B.2)$$
and
\[ t^- = T^- + \frac{D(T^-)}{s} \]  \hspace{1cm} (B.3)

which becomes
\[ t^- = \frac{\alpha}{\alpha + \gamma} T^- - \frac{\beta}{\alpha - \beta} t_q + \frac{\alpha(\beta + \gamma)}{(\alpha - \beta)(\alpha + \gamma)} \tilde{t} - \frac{\tau}{\alpha - \beta} \]  \hspace{1cm} (B.4)
\[ \frac{\alpha}{\alpha + \gamma} T^- = t^- + \frac{\beta}{\alpha - \beta} t_q - \frac{\alpha(\beta + \gamma)}{(\alpha - \beta)(\alpha + \gamma)} \tilde{t} + \frac{\tau}{\alpha - \beta} \]  \hspace{1cm} (B.5)
\[ T^- = \frac{\alpha + \gamma}{\alpha} t^- + \frac{\beta(\alpha + \gamma)}{\alpha(\alpha - \beta)} t_q - \frac{\beta + \gamma}{\alpha - \beta} \tilde{t} + \frac{\alpha + \gamma}{\alpha(\alpha - \beta)} \tau \]  \hspace{1cm} (B.6)

Assuming that \( \tilde{t} \in [T^+ + \tau/\alpha, T^-] \), the total number of commuters \( N \) is given by
\[ N = r_1 [T^+ - t_q + \tilde{t} - (T^+ + \tau/\alpha)] + r_2 [T^- - \tilde{t}] + \frac{2s\tau}{\alpha + \gamma} \]
\[ = \frac{s\alpha}{\alpha - \beta} (\tilde{t} - t_q - \tau/\alpha) + \frac{s\alpha}{\alpha + \gamma} [T^- - \tilde{t}] + \frac{2s\tau}{\alpha + \gamma} \]  \hspace{1cm} (B.7)

\[ \frac{N(\alpha - \beta)(\alpha + \gamma)}{s\alpha} = (\alpha + \gamma)(\tilde{t} - t_q - \tau/\alpha) + (\alpha - \beta)(T^- - \tilde{t}) + \frac{(\alpha - \beta)2\tau}{\alpha} \]
\[ = \tilde{t}((\alpha + \gamma) - (\alpha - \beta)) - t_q(\alpha + \gamma) + T^- (\alpha - \beta) + \tau \left( \frac{2(\alpha - \beta)}{\alpha} - \frac{\alpha + \gamma}{\alpha} \right) \]
\[ = \tilde{t}(\beta + \gamma) - t_q(\alpha + \gamma) + T^- (\alpha - \beta) + \tau \left( \frac{\alpha - 2\beta - \gamma}{\alpha} \right) \]  \hspace{1cm} (B.8)

So, for equilibrium between \( t_q \) and \( \tilde{t} \) we have
\[ \beta(t^* - t_q) = \alpha(t^* - \tilde{t}) + \tau \]  \hspace{1cm} (B.9)

Now from the definition of \( \tilde{t} \) it follows that
\[ t^* = \tilde{t} + D(\tilde{t})/s \]
\[ = \tilde{t} + \frac{s}{\alpha - \beta} \left( T^+ - t_q + \tilde{t} - (T^+ + \tau/\alpha) \right) - s(\tilde{t} - t_q) \]
\[ = \tilde{t} - \frac{\alpha}{\alpha - \beta} t_q - \frac{\beta}{\alpha - \beta} - \frac{\tau}{\alpha - \beta} \]  \hspace{1cm} (B.10)

that is
\[ t_q = \frac{\alpha}{\beta} t - \frac{\alpha - \beta}{\beta} t^* - \frac{\tau}{\beta} \]  \hspace{1cm} (B.11) 

and

\[ \ddot{i} = \frac{\beta}{\alpha} t_q + \frac{\alpha - \beta}{\alpha} t^* + \frac{\tau}{\alpha} \]  \hspace{1cm} (B.12)

Combining (B.8) and (B.9) by substituting for \( \ddot{i} \) in (B.8) yields

\[ \frac{N(\alpha - \beta)(\alpha + \gamma)}{s\alpha} = (\beta t_q + [\alpha - \beta] t^* + \tau) \frac{\beta + \gamma}{\alpha} - t_q(\alpha + \gamma) \]

\[ + T^- (\alpha - \beta) + \tau \frac{\alpha - 2\beta - \gamma}{\alpha} \]

\[ = t^*(\alpha - \beta)(\beta + \gamma) - t_q[\beta(\beta + \gamma) - (\alpha + \gamma)] + T^- (\alpha - \beta) + \tau(\frac{\alpha - 2\beta - \gamma + \beta + \gamma}{\alpha}) \]

\[ = t^*(\alpha - \beta)(\beta + \gamma) + t_q[\beta^2 + \beta \gamma - \alpha^2 - \alpha \gamma] + T^- (\alpha - \beta) + \tau(\alpha - \beta) \]

\[ = t^*(\alpha - \beta)(\beta + \gamma) - t_q[\beta(\alpha + \gamma) - (\alpha + \gamma)] + T^- (\alpha - \beta) + \tau(\alpha - \beta) \]

\[ = t^*(\alpha - \beta)(\beta + \gamma) - t_q(\beta + \alpha + \gamma) + T^- (\alpha - \beta) + \tau \]  \hspace{1cm} (B.13)

\[ t_q = \frac{t^*(\beta + \gamma)}{\beta + \alpha + \gamma} + \frac{\alpha}{\beta + \alpha + \gamma} T^- - \frac{N(\alpha + \gamma)}{s(\beta + \alpha + \gamma)} + \frac{\tau}{\beta + \alpha + \gamma} \]  \hspace{1cm} (B.14)

Further substituting \( \ddot{i} \) from (B.12) into (B.6) yields

\[ T^- = \frac{\alpha + \gamma}{\alpha} t^- + \frac{\beta(\alpha + \gamma)}{\alpha(\alpha - \beta)} t_q - \frac{\beta + \gamma}{\alpha - \beta} \frac{\beta}{\alpha} t^* + \frac{\alpha - \beta}{\alpha} t^* + \frac{\alpha + \gamma}{\alpha} t^- \]  \hspace{1cm} (B.15)

\[ T^- = \frac{\alpha + \gamma}{\alpha} t^- + \frac{\beta}{\alpha} t_q - \frac{\beta + \gamma}{\alpha} t^* + \frac{\tau}{\alpha} \]  \hspace{1cm} (B.16)

and substituting for \( T^- \) in (B.14) yields

\[ t_q = t^- + \frac{2\tau}{\alpha + \gamma} - \frac{N}{s} \]  \hspace{1cm} (B.17)

\[ \ddot{i} = \frac{\beta}{\alpha m} (t^*(\beta + \gamma) + \alpha T^- - \frac{N(\alpha + \gamma)}{s} + \tau) + \frac{\alpha - \beta}{\alpha} t^* + \frac{\tau}{\alpha} \]

\[ = (t^*(\alpha + \gamma) + \beta T^- - \frac{N\beta(\alpha + \gamma)}{s\alpha} + \tau \frac{\alpha + 2\beta + \gamma}{\alpha}) \]  \hspace{1cm} (B.18)
where \( m = \alpha + \beta + \gamma \).

Note that we are assuming that the queue length is never zero for a positive amount of time. This assumption implies that

\[
D(T^+ + \tau/\alpha) \geq 0
\]  

which implies

\[
\frac{s\alpha}{\alpha - \beta} (T^+ - t_q) - s(T^+ + \tau/\alpha - t_q) \geq 0
\]  

\[
T^+ \geq t_q + \frac{\tau(\alpha - \beta)}{\alpha \beta}
\]  

Note, for further use below, that the length of the period during which commuters depart at rate \( r_1 \) is given by \( (\hat{t} - t_q - \tau/\alpha) \).

Similarly, the queue should not be of zero length after the toll begins during the toll period; in other words, \( D(T^-) \geq 0 \),

\[
D(T^-) \geq 0
\]  

\[
\frac{s\alpha}{\alpha - \beta} (\hat{t} - t_q - \tau/\alpha) + \frac{s\alpha}{\alpha + \gamma} (T^- - \hat{t}) - s(T^- - t_q) \geq 0
\]  

\[
\frac{\gamma}{\alpha + \gamma} T^- \leq \frac{\alpha(\beta + \gamma)}{(\alpha - \beta)(\alpha + \gamma)} \hat{t} - \frac{\beta}{\alpha - \beta} t_q - \frac{\tau}{\alpha - \beta}
\]  

\[
T^- \leq \frac{\alpha(\beta + \gamma)}{\gamma(\alpha - \beta)} \hat{t} - \frac{\beta(\alpha + \gamma)}{\gamma(\alpha - \beta)} t_q - \frac{\tau(\alpha + \gamma)}{\gamma(\alpha - \beta)}
\]

**B.2 The Optimal Step-Toll**

To derive the optimal step-toll it should be noted that the total cost, excluding tolls, which we need to minimize is given by \( TC \)

\[
TC = \beta(t^* - t_q)N - (t^- - t^+)s\tau
\]  

Substituting for \( t^+ \), and \( t^- \) and simplifying

\[
TC = \beta(t^* - t_q)N - (\frac{\alpha}{\alpha + \gamma} T^- - \frac{\tau}{\alpha - \beta} + \frac{\alpha(\beta + \gamma)}{(\alpha - \beta)(\alpha + \gamma)} \hat{t} - \frac{\alpha}{\alpha - \beta} T^+)s\tau
\]  

Substituting for \( t_q \) and \( \hat{t} \)

\[
TC = \beta(t^* - (t^*(\frac{\beta + \gamma}{m} + \frac{\alpha}{m} T^- - \frac{N(\alpha + \gamma)}{sm} + \frac{\tau}{m}))N - (\frac{\alpha}{\alpha + \gamma} T^- - \frac{\tau}{\alpha - \beta})
\]
\[ T^+ \geq t_q + \frac{\tau (\alpha - \beta)}{\alpha \beta} \]  
(B.29)

\[ T^+ - t_q - \frac{\tau (\alpha - \beta)}{\alpha \beta} \geq 0 \]  
(B.30)

\[ -T^+ + t_q + \frac{\tau (\alpha - \beta)}{\alpha \beta} \leq 0 \]  
(B.31)

and

\[ T^- \leq \frac{\alpha (\beta + \gamma)}{\gamma (\alpha - \beta)} \cdot \frac{\beta (\alpha + \gamma)}{\gamma (\alpha - \beta)} \cdot t_q - \frac{\tau (\alpha + \gamma)}{\gamma (\alpha - \beta)} \]  
(B.32)

\[ T^- - \frac{\alpha (\beta + \gamma)}{\gamma (\alpha - \beta)} \left( t^* \cdot \frac{\alpha + \gamma}{m} + \frac{\beta}{m} T^- - \frac{N (\alpha + \gamma)}{s m} + \frac{\tau}{m} \right) + \frac{\tau (\alpha + \gamma)}{\gamma (\alpha - \beta)} \leq 0 \]  
(B.33)

Substituting \( t_q \) and \( \bar{t} \) in (B.31) and (B.33)

\[ -T^+ + \left( t^* \frac{(\beta + \gamma)}{m} + \frac{\alpha}{m} T^- - \frac{N (\alpha + \gamma)}{s m} + \frac{\tau}{m} \right) + \frac{\tau (\alpha - \beta)}{\alpha \beta} \leq 0 \]  
(B.34)

\[ T^- - \frac{\alpha (\beta + \gamma)}{\gamma (\alpha - \beta)} \left( t^* \cdot \frac{\alpha + \gamma}{m} + \frac{\beta}{m} T^- - \frac{N (\alpha + \gamma)}{s m} + \frac{\tau}{m} \right) + \frac{\tau (\alpha + \gamma)}{\gamma (\alpha - \beta)} \leq 0 \]  
(B.35)

We can now set up the KKT conditions. Thus

\[ TC = \beta (t^* - \left( t^* \frac{(\beta + \gamma)}{m} + \frac{\alpha}{m} T^- - \frac{N (\alpha + \gamma)}{s m} + \frac{\tau}{m} \right)) + \frac{T^- - \tau}{(\beta + \alpha + \gamma)} \cdot N - \left[ \frac{\alpha}{\alpha + \gamma} - \frac{\tau}{(\beta + \alpha + \gamma)} \right] \cdot N \]

\[ + \frac{\alpha (\beta + \gamma)}{(\alpha - \beta) (\alpha + \gamma)} \left( t^* \cdot \frac{\alpha + \gamma}{m} + \frac{\beta}{m} T^- - \frac{N (\alpha + \gamma)}{s m} + \frac{\tau}{m} \right) + \frac{\tau (\alpha - \beta)}{\alpha \beta} \cdot s \tau \]

\[ + \lambda_1 (-T^+ + \left( t^* \frac{(\beta + \gamma)}{m} + \frac{\alpha}{m} T^- - \frac{N (\alpha + \gamma)}{s m} + \frac{\tau}{m} \right) + \frac{\tau (\alpha - \beta)}{\alpha \beta} \]

\[ + \lambda_2 (T^- - \frac{\alpha (\beta + \gamma)}{\gamma (\alpha - \beta)} \left( t^* \cdot \frac{\alpha + \gamma}{m} + \frac{\beta}{m} T^- - \frac{N (\alpha + \gamma)}{s m} + \frac{\tau}{m} \right) + \frac{\tau (\alpha + \gamma)}{\gamma (\alpha - \beta)} \)

\[ + \frac{(\beta + \gamma)}{\gamma (\alpha - \beta)} \left( t^* \cdot \frac{\alpha + \gamma}{m} + \frac{\alpha}{m} T^- - \frac{N (\alpha + \gamma)}{s m} + \frac{\tau}{m} \right) + \frac{\tau (\alpha + \gamma)}{\gamma (\alpha - \beta)} \]  
(B.37)
Letting the Lagrangian be denoted by $KT$,

$$
\frac{\partial KT}{\partial \tau} = -\beta N - s\left[\frac{ma}{\alpha + \gamma}T - \frac{2m\tau}{\alpha - \beta} + \frac{\alpha(\beta + \gamma)}{\alpha - \beta}t^* + \frac{\alpha\beta(\beta + \gamma)}{(\alpha - \beta)(\alpha + \gamma)}T^-\right] \\
- \frac{N\beta(\beta + \gamma)}{s(\alpha - \beta)} + \frac{2\tau(\alpha + 2\beta + \gamma)(\beta + \gamma)}{(\alpha - \beta)(\alpha + \gamma)} - \frac{ma}{\alpha - \beta}T^+ + \lambda_1\left[1 + \frac{m(\alpha - \beta)}{\alpha\beta}\right] \\
+ \lambda_2\left[-\frac{\alpha(\beta + \gamma)(\alpha + 2\beta + \gamma)}{\gamma(\alpha - \beta)} + \frac{\beta(\alpha + \gamma)}{\gamma(\alpha - \beta)} + \frac{m(\alpha + \gamma)}{\gamma(\alpha - \beta)}\right] = 0
$$

(B.38)

which implies

$$
\frac{\partial KT}{\partial \tau} = N\frac{\beta(2\beta - \alpha + \gamma)}{\alpha - \beta} - \frac{s\alpha^2}{\alpha - \beta}T - s\frac{\alpha(\beta + \gamma)}{\alpha - \beta}t^* - \frac{2\tau s[2\beta(\beta + \gamma) - \alpha(\alpha + \gamma)]}{(\alpha - \beta)(\alpha + \gamma)} \\
+ \frac{ma}{\alpha - \beta}T^+ + \lambda_1\frac{\alpha\beta + m(\alpha - \beta)}{\alpha\beta} + \lambda_2\frac{\alpha + 2\beta + \gamma}{\gamma} = 0
$$

(B.39)

$$
\frac{\partial KT}{\partial T^-} = -\alpha\beta - \frac{ma}{\alpha + \gamma} + \frac{\alpha\beta(\beta + \gamma)}{(\alpha - \beta)(\alpha + \gamma)}T^+ + \lambda_1\alpha \\
+ \lambda_2[m - \frac{\alpha\beta(\beta + \gamma)}{\gamma(\alpha - \beta)} + \frac{\alpha\beta(\alpha + \gamma)}{\gamma(\alpha - \beta)}] = 0
$$

(B.40)

which further implies

$$
-\alpha\beta N - s\tau\left[\frac{ma}{\alpha + \gamma} + \frac{\alpha\beta(\beta + \gamma)}{(\alpha - \beta)(\alpha + \gamma)}\right] + \lambda_1\alpha + \lambda_2\left[\frac{(\alpha + \gamma)(\beta + \gamma)}{\gamma}\right] = 0
$$

$$
-\alpha\beta N - \frac{s\tau\alpha^2}{(\alpha - \beta)} + \lambda_1\alpha + \lambda_2\left[\frac{(\alpha + \gamma)(\beta + \gamma)}{\gamma}\right] = 0
$$

$$
-\beta N - \frac{s\tau\alpha}{(\alpha - \beta)} + \lambda_1 + \lambda_2\left[\frac{(\alpha + \gamma)(\beta + \gamma)}{\gamma}\right] = 0
$$

(B.41)

$$
\frac{\partial KT}{\partial T^+} = \frac{\alpha}{\alpha - \beta}s\tau - \lambda_1 = 0
$$

(B.42)

or

$$
\frac{\alpha}{\alpha - \beta}s\tau = \lambda_1
$$

(B.43)

The two constraints yield

$$
\lambda_1[-(-T^+ + (t^*\frac{(\beta + \gamma)}{m} + \frac{\alpha}{m}T - \frac{N(\alpha + \gamma)}{sm} + \frac{\tau}{m}) + \tau(\alpha - \beta)) + \frac{\alpha(\beta + \gamma)}{\alpha\beta}T^-] = 0
$$
\[ \lambda_1[-T^+ + \left( t^* \frac{(\beta + \gamma)}{m} + \frac{\alpha}{m} T^- - \frac{N(\alpha + \gamma)}{sm} + \frac{\tau}{m} + \frac{\tau(\alpha - \beta)}{\alpha \beta} \right) = 0 \quad (B.44) \]

\[ \lambda_2[-(T^- - \frac{\alpha(\beta + \gamma)}{\gamma(\alpha - \beta)}(t^* \frac{\alpha + \gamma}{m} + \frac{\beta}{m} T^- - \frac{N\beta(\alpha + \gamma)}{s\alpha(m)} + \frac{\tau}{m} + \frac{\tau(\alpha + \gamma)}{\gamma(\alpha - \beta)})] = 0 \quad (B.45) \]

\[ \lambda_1, \lambda_2 \geq 0 \quad (B.46) \]

Assume \( \lambda_1 > 0, \lambda_2 > 0 \). Then we can substitute for \( \lambda_1 \) in (B.41) which yields

\[ \lambda_2 = \frac{\alpha \beta \gamma N}{(\alpha + \gamma)(\beta + \gamma)} \quad (B.47) \]

It follows from (B.44) that

\[ T^+ = \left( t^* \frac{(\beta + \gamma)}{m} + \frac{\alpha}{m} T^- - \frac{N(\alpha + \gamma)}{sm} + \frac{\alpha \beta + m(\alpha - \beta)}{\alpha \beta m} \right) \quad (B.48) \]

and from (B.2), (B.45) that

\[ t^* \left[ \frac{\beta(\alpha + \gamma)(\beta + \gamma)}{\gamma(\alpha - \beta)m} - \frac{\alpha(\alpha + \gamma)(\beta + \gamma)}{\gamma(\alpha - \beta)m} + T^- \left[ \frac{\alpha \beta(\alpha + \gamma)}{\gamma(\alpha - \beta)m} - \frac{\alpha \beta(\beta + \gamma)}{\gamma(\alpha - \beta)m} + 1 \right] \right. \]
\[ \left. + \frac{N}{sm \gamma(\alpha - \beta)}[\beta(\alpha + \gamma)(\beta + \gamma) - \beta(\alpha + \gamma)(\alpha + \gamma)] \quad (B.49) \right. \]
\[ + \frac{\tau}{\gamma(\alpha - \beta)m} \left[ -\frac{\alpha(\beta + \gamma)(\alpha + 2\beta + \gamma)}{\alpha} + \beta(\alpha + \gamma) + (\alpha + \gamma)m \right] = 0 \quad (B.50) \]

This reduces to

\[ -t^*(\alpha + \gamma)(\beta + \gamma) + T^- (\alpha \beta + \gamma m) - N \frac{\beta(\alpha + \gamma)}{s} + \tau(\alpha + 2\beta + \gamma) = 0 \quad (B.51) \]

which yields

\[ -t^*(\alpha + \gamma)(\beta + \gamma) + T^- (\alpha + \gamma)(\beta + \gamma) - N \frac{\beta(\alpha + \gamma)}{s} + \tau(\alpha + 2\beta + \gamma) = 0 \]

\[ T^- = \left( \frac{1}{(\alpha + \gamma)(\beta + \gamma)} \left[ t^*(\alpha + \gamma)(\beta + \gamma) + N \frac{\beta(\alpha + \gamma)}{s} - \tau(\alpha + 2\beta + \gamma) \right] \right) \]

or

\[ T^- = t^* + N \frac{\beta}{s(\beta + \gamma)} - \tau \frac{(\alpha + 2\beta + \gamma)}{(\alpha + \gamma)(\beta + \gamma)} \quad (B.53) \]

Substituting \( T^+ \) from (B.48) in (B.2), (B.39)
Appendix B  Deriving the Optimal Toll

\[
\begin{align*}
\frac{\partial KT}{\partial \tau} &= N\beta(2\beta - \alpha + \gamma) - \frac{sa^2}{\alpha - \beta}T - \frac{\alpha(\beta + \gamma)}{\alpha - \beta}t^* - 2\tau s[2\beta(\beta + \gamma) - \alpha(\alpha + \gamma)] + \frac{mas}{\alpha - \beta}(t^*(\beta + \gamma) + \frac{\alpha}{m}T) - \frac{N(\alpha + \gamma)}{sm} \quad (B.54) \\
+ \frac{\alpha \beta + m(\alpha - \beta)}{\alpha \beta m} + \lambda_1 \frac{\alpha \beta + m(\alpha - \beta)}{\alpha \beta} + \lambda_2 \frac{\alpha + 2\beta + \gamma}{\gamma} &= 0
\end{align*}
\]

Substituting \(\lambda_1, \lambda_2\) in the above

\[
\begin{align*}
\frac{\partial KT}{\partial \tau} &= N\beta(2\beta - \alpha + \gamma) - \frac{sa^2}{\alpha - \beta}T - \frac{\alpha(\beta + \gamma)}{\alpha - \beta}t^* - 2\tau s[2\beta(\beta + \gamma) - \alpha(\alpha + \gamma)] \\
+ \frac{\alpha s}{\alpha - \beta}(t^*(\beta + \gamma) + \alpha T) - \frac{N(\alpha + \gamma)}{s} + \tau \frac{\alpha \beta + m(\alpha - \beta)}{\alpha \beta} \\
+ \frac{\alpha}{\alpha - \beta} \frac{\alpha \beta + m(\alpha - \beta)}{\alpha \beta} + \frac{\alpha \beta \gamma N}{(\alpha + \gamma)(\beta + \gamma)} \frac{\alpha + 2\beta + \gamma}{\gamma} &= 0 \quad (B.55)
\end{align*}
\]

Simplifying

\[
\begin{align*}
\frac{\partial KT}{\partial \tau} &= N\beta(2\beta - \alpha + \gamma) - \frac{2\tau s[2\beta(\beta + \gamma) - \alpha(\alpha + \gamma)]}{(\alpha - \beta)(\alpha + \gamma)} \\
- \frac{\alpha s}{\alpha - \beta} \frac{N(\alpha + \gamma)}{s} + 2\tau \frac{\alpha \beta + m(\alpha - \beta)}{\beta(\alpha - \beta)} + \frac{\alpha \beta N}{(\alpha + \gamma)(\beta + \gamma)}(\alpha + 2\beta + \gamma) &= 0 \quad (B.56) \\
N\beta(2\beta - \alpha + \gamma) - \frac{2\tau s[2\beta(\beta + \gamma) - \alpha(\alpha + \gamma)]}{(\alpha - \beta)(\alpha + \gamma)} + \frac{\alpha s}{\alpha - \beta} \left( - \frac{N(\alpha + \gamma)}{s} \right) \\
+ 2\tau \frac{\alpha \beta + m(\alpha - \beta)}{\beta(\alpha - \beta)} + \frac{\alpha \beta N}{(\alpha + \gamma)(\beta + \gamma)}(\alpha + 2\beta + \gamma) &= 0 \quad (B.57) \\
2\tau \frac{\beta(2\beta(\beta + \gamma) - \alpha(\alpha + \gamma)) - \alpha \beta \gamma N}{\beta(\alpha - \beta)(\alpha + \gamma)} \frac{\alpha + 2\beta + \gamma}{\gamma} &= N\left[ \frac{\beta(2\beta - \alpha + \gamma) - \alpha(\alpha + \gamma)}{(\alpha + \gamma)}(\beta + \gamma)(\alpha - \beta) \right] \frac{\alpha + 2\beta + \gamma}{\gamma} \quad (B.58)
\end{align*}
\]

\[
\tau = \frac{N}{s} \beta \frac{\beta(2\beta(\beta + \gamma) - \alpha(\alpha + \gamma))(\alpha + \gamma)(\beta + \gamma) + \alpha \beta(\alpha + 2\beta + \gamma)(\alpha - \beta)}{2(\beta + \gamma)} - \frac{\beta [2\beta(\beta + \gamma) - \alpha(\alpha + \gamma)] - [\alpha \beta + m(\alpha - \beta)](\alpha + \gamma)}{2(\beta + \gamma)} \\
= \frac{N}{s} \beta \gamma \frac{\beta \gamma^2 + \alpha \beta^2 - \alpha^3 + 2\beta^3 - \alpha \gamma^2 - 2\alpha^2 \beta + 3\beta^2 \gamma - 2\alpha^2 \gamma - \alpha \beta \gamma}{2(\beta + \gamma)} - \frac{\beta [2\beta(\beta + \gamma) - \alpha(\alpha + \gamma)] - [\alpha \beta + m(\alpha - \beta)](\alpha + \gamma)}{2(\beta + \gamma)} \\
= \frac{N}{s} \beta \gamma \frac{\beta \gamma^2 + \alpha \beta^2 - \alpha^3 + 2\beta^3 - \alpha \gamma^2 - 2\alpha^2 \beta + 3\beta^2 \gamma - 2\alpha^2 \gamma - \alpha \beta \gamma}{2(\beta + \gamma)} \\
= \frac{N}{s} \beta \gamma \frac{2\beta \gamma^3 + 2\beta^3 \gamma - \alpha^3 - 2\alpha^2 \gamma - 2\alpha^2 \beta - \alpha \beta \gamma - \alpha \gamma^2 - 2\beta \gamma^2 + \alpha \beta^2}{2(\beta + \gamma)} \\
\tau &= \frac{N}{s} \beta \gamma \quad (B.59)
\]

174
Appendix C

Correcting the Uniform Toll

We demonstrate that the uniform portion of the optimal step toll reported by Arnott et al. (1990b) is incorrect. We also derive the correct expression for this toll.

We use the notation introduced in Arnott et al. (1990b). In particular, $\alpha$ denotes the dollar cost of travel time, $\beta$ denotes the dollar cost of early arrival time, $\gamma$ denotes the dollar cost of late arrival time, $N^*_i$ denotes the number of commuters on route $i$ in the presence of a step toll, $N$ denotes the total number of commuters, $s_i$ denotes the service rate of the deterministic queue on route $i$, $T^f_i$ denotes the free-flow travel time on route $i$, $C^*_i(t)$ denotes the user cost of travel on route $i$ at time $t$ in the presence of a step toll, $C^*_i$ denotes the equilibrium user cost of travel on route $i$, $t^*$ denotes the (common) desired arrival time at the destination, $t^*_{0,i}$ denotes the time of the first departure on route $i$ in the presence of a step toll, $\Delta$ denotes one half of the acceptable arrival time window around $t^*$, $\tau^*_i$ denotes the value of the time-varying portion of the step toll on route $i$, and $\pi^*$ denotes the uniform portion of the step toll (on route 1).

C.1 The Error in the Reported Results

Perhaps the easiest way to see that there is an error in the results reported by Arnott et al. (1990b) is to consider the numerical example they present in Table 3. First, observe that since all users experience the same cost in equilibrium, the equilibrium user costs on the two routes are given by

\begin{align}
C^*_1 &= \alpha T^f_1 + \beta[(t^* - \Delta) - (t^*_{q,1} + T^f_1)] + \pi^* \\
C^*_2 &= \alpha T^f_2 + \beta[(t^* - \Delta) - (t^*_{q,2} + T^f_2)]
\end{align}

which is simply the cost to the commuter that departs just as the queue begins to form.
Now, letting
\[ \delta = \frac{\beta \gamma}{\beta + \gamma} \]  
(C.3)
we know from their (33b) and (33a) that
\[ t_{q,i}^* = t^* - \Delta - T_i^f - \frac{\gamma}{\beta + \gamma} \left( \frac{N_i^s}{s_i} - 2\Delta \right) + \frac{(\gamma - \alpha)\tau_i^s}{(\beta + \gamma)(\alpha + \gamma)} \]  
(C.4)
\[ \tau_i^* = \delta \frac{N_i^s}{2s_i} \]  
(C.5)
for \( i = 1, 2 \). It thus follows that
\[ C_2^* = \alpha T_2^f + \beta \left[ \frac{-2\gamma \Delta}{\beta + \gamma} + \frac{N_2^s}{s_2} \left( \frac{\gamma}{\beta + \gamma} - \delta \frac{\gamma - \alpha}{2(\beta + \gamma)(\alpha + \gamma)} \right) \right] \]  
(C.6)
\[ = \alpha T_2^f + \beta \left[ \frac{-2\gamma \Delta}{\beta + \gamma} + \frac{N_2^s}{s_2} \left( \delta - \delta \frac{\beta(\gamma - \alpha)}{2(\beta + \gamma)(\alpha + \gamma)} \right) \right] \]
\[ = \alpha T_2^f - \frac{2\beta \gamma \Delta}{\beta + \gamma} + \frac{N_2^s}{s_2} \left( \delta - \delta \frac{\beta(\gamma - \alpha)}{2(\beta + \gamma)(\alpha + \gamma)} \right) \]
\[ = \alpha T_2^f - 2\delta \Delta + \delta \frac{N_2^s}{s_2} \left( 1 - \frac{\beta(\gamma - \alpha)}{2(\beta + \gamma)(\alpha + \gamma)} \right). \]
So, letting \( B = \left[ 1 - \frac{\beta(\gamma - \alpha)}{2(\beta + \gamma)(\alpha + \gamma)} \right] = \frac{2(\beta + \gamma)(\alpha + \gamma) - \beta(\gamma - \alpha)}{2(\beta + \gamma)(\alpha + \gamma)} \) it further follows that
\[ C_2^* = \alpha T_2^f - 2\delta \Delta + B\delta \frac{N_2^s}{s_2} \]  
(C.7)
and similarly that
\[ C_1^* = \alpha T_1^f - 2\delta \Delta + B\delta \frac{N_1^s}{s_1} + \pi^*. \]  
(C.8)
If we now substitute the values given in Table 3 we see that \( C_1^* = 6.11 \) and \( C_2^* = 6.01 \), which is clearly not an equilibrium. Hence, there must be an error in the results they present.

### C.2 Deriving the Optimal Toll

As it turns out, the mistake is in the unnumbered equation on page 220 [between their (34c) and (35)]. It is claimed that
C.2 Deriving the Optimal Toll

\[ C_1^* = C_2^* \Rightarrow \alpha T_1^f + \delta \left[ \frac{A N_1^f}{2 s_1} - 2\Delta \right] + \pi^* = \alpha T_2^f + \delta \left[ \frac{A N_2^f}{2 s_2} - 2\Delta \right] \]  \hspace{1cm} (C.9)

where \( A = \frac{3(\beta + \gamma)(\alpha + \gamma) - \beta(\gamma - \alpha)}{2(\beta + \gamma)(\alpha + \gamma)} \). However, it is easy to see from the discussion above that

\[ C_1^* = C_2^* \Rightarrow \alpha T_1^f - 2\delta \Delta + B\delta \frac{N_1^f}{s_1} + \pi^* = \alpha T_2^f - 2\delta \Delta + B\delta \frac{N_2^f}{s_2} \]  \hspace{1cm} (C.10)

and hence that

\[ \pi^* = \alpha(T_2^f - T_1^f) + B\delta \left( \frac{N_2^f}{s_2} - \frac{N_1^f}{s_1} \right). \]  \hspace{1cm} (C.11)

Now, since we know from their (34a) that

\[ N_i^* = \frac{s_1 s_2}{s_1 + s_2} \left[ \frac{N}{s_j} + \frac{\alpha}{\delta A} (T_j^f - T_i^f) \right] \hspace{1cm} j \neq i \]  \hspace{1cm} (C.12)

for \( i = 1, 2 \) it follows that

\[ \frac{N_2^f}{s_2} - \frac{N_1^f}{s_1} = \frac{s_1}{s_1 + s_2} \left[ \frac{N}{s_1} + \frac{\alpha}{\delta A} (T_1^f - T_2^f) \right] - \frac{s_2}{s_1 + s_2} \left[ \frac{N}{s_2} + \frac{\alpha}{\delta A} (T_2^f - T_1^f) \right] \]  \hspace{1cm} (C.13)

\[ = \frac{s_1}{s_1 + s_2} \frac{\alpha}{\delta A} (T_1^f - T_2^f) - \frac{s_2}{s_1 + s_2} \frac{\alpha}{\delta A} (T_2^f - T_1^f) \]  \hspace{1cm} (C.14)

\[ = \frac{s_1 \alpha(T_1^f - T_2^f)}{(s_1 + s_2) \delta A} \]  \hspace{1cm} (C.15)

Thus, substituting (C.13) into (C.11) we have

\[ \pi^* = \alpha(T_2^f - T_1^f) + \frac{B A}{\alpha} \alpha(T_1^f - T_2^f) \]  \hspace{1cm} (C.15)

\[ = \alpha \left( 1 - \frac{B}{A} \right) (T_2^f - T_1^f) \]  \hspace{1cm} (C.15)

\[ = \alpha \left( 1 - \frac{2(\beta + \gamma)(\alpha + \gamma) - \beta(\gamma - \alpha)}{3(\beta + \gamma)(\alpha + \gamma) - \beta(\gamma - \alpha)} \right) (T_2^f - T_1^f) \]  \hspace{1cm} (C.15)

\[ = \frac{\alpha(\beta + \gamma)(\alpha + \gamma)}{3(\beta + \gamma)(\alpha + \gamma) - \beta(\gamma - \alpha)} (T_2^f - T_1^f). \]  \hspace{1cm} (C.15)
Returning to the example in Table 3, we find that the correct value for $\pi^*$ is $0.22$ instead of $0.32$. With this toll in place, users of both routes experience costs (including all tolls) of $6.01$ (which is $0.10$ less than they would experience in the no-toll equilibrium).
Appendix D

How Information Results in Benefits

In this section we attempt to answer the following question: How do net benefits come about using information when capacity is not directly increased?

D.1 Motivation

The previous chapters showed that net benefits, in terms of total costs, always accrued when information was supplied to an unpriced regime. The general motivation for designing information systems is well-known: Provide information to improve utilization of the network, and travel times will go down. Whatever the merits of this argument, it appears that a priori we can foresee, for real-world networks, only two impacts of an accurate information system: The total system costs for those travelers who as a result of the information provision are shifted from the routes that suffered a capacity loss due to an incident to incident free routes, will go down. Similarly, the costs for those commuters who travel on the incident-free routes will go up as a result of the diverted travelers. As capacity is not increased as a result of providing information, the gains for the diverted group must outweigh the losses suffered by the undiverted group if there are to be net positive benefits.

Since it is not intuitively obvious why this should occur, both on real-world networks and for our model, we outline an analysis for our specific model that explains when net benefits may occur. Note that the problems associated with information systems, such as concentration and overreaction, have been attributed to imperfect prediction. Work of course continues to develop better prediction models. Since our information system is assumed perfect in this regard, the following analysis can help to determine an upper bound
on possible gains for myopic systems\(^1\).

As will become clear during the analysis, we need to look at two cases, one where the departure rate is greater on the incident route and one where it is lower. The results presented here are based on the network used elsewhere in this thesis (see Figure 4-1). Since the objective is primarily to explore some relationships between benefits due to information, simplifying assumptions are made whenever appropriate to extract the essentials. Attention is restricted to the no-toll regime. We assume that the capacities or service rates of links 1 and 2 can vary, and denote them by \(s_1\) and \(s_2\) respectively. Furthermore, without loss of generality, the incident is assumed to occur on link 1 (i.e. route 1). The capacity during the incident is given by \(s_a < s_1\). Denote the time of the beginning and of the end of the incident by \(t_a\) and \(t_e\) respectively.

**D.2 Analyzing Benefits: \(s_1 < s_2\).**

First observe that there is a certain number \(N_T\) of commuters who are trapped between \(B\) and \(C\) at the moment the incident occurs, and cannot switch routes. Then \(N_T\) is equal to the number of commuters who have passed \(B\) at time \(t_a\), minus the number of commuters who have exited the bottleneck on route 1 (i.e. have arrived) which can be written as

\[
N_T = \int_{t_1^a}^{t_a-t_{AB}} r_1(t) \, dt - (tb - t_1^C) s_1
\]

(D.1)

where \(t_{AB}\) is the travel time between \(A\) and \(B\), \(t_1^a\) is the time of the first departure on route 1, and \(t_1^C\) is the time of the first commuter to arrive at the bottleneck on route 1.

We shall be deriving benefits in the manner of the benefitograms, such that we look at one commuter at a time, and compare his costs under information to those in the no information scenario. In addition “packets” are analyzed, which is a conveniently chosen number of travelers on both routes departing during a short interval at the point in time in question.

**Benefits for First Packet**

Thus, we denote the no-information costs for a commuter arriving at node \(B\) at time \(t\) with default route \(i\) by \(c_i^n(t)\). For the information case, the corresponding term is \(C_i(t)\). Finally, \(C_{EQ}\) denotes the equilibrium cost.

The total number \(N_A\) of travelers who pass through the bottleneck at its reduced capacity in the no-information case is given by

\(^1\)Of course, since myopic systems do not aim at minimizing system costs, it is possible to have greater benefits with an imperfect system. Being a best-case analysis, the above helps to show where the development of better predictive capabilities for myopic systems leads.
\[ N_A = s_a d \]  \hspace{1cm} (D.2)

where \( d = t_e - t_o \) denotes the duration of the incident. Of these, the number of travelers who are ahead of the commuter who is at \( B \) at time \( t_a \) (the number of "incident-trapped" commuters) is given by \( N'_A = \min(N_A, N_T) \).

As the information system begins operation, the first commuter whose default route is 1 to receive guidance at \( B \), at time \( t_a \), will have,

\[ C^N_1(t_a) = C_{EQ} + \omega(N'_A \ast \Delta s) \]  \hspace{1cm} (D.3)

where \( \Delta s = \text{increase in service time per traveler due to incident} = 1/s_a - 1/s_1 \), and \( \omega = \alpha + \gamma \). Throughout, we are assuming for simplicity that this commuter would have arrived after \( t^* \) even without the occurrence of the incident. Moreover,

\[ C^Y_1(t_a) = C_{EQ} \]  \hspace{1cm} (D.4)

assuming he merges smoothly into the stream of travelers on route 2, incurring equilibrium cost. Then his benefits are given by

\[ \omega(C^Y_1(t_a) - C^N_1(t_a)) = \omega N'_A \ast \Delta s \]  \hspace{1cm} (D.5)

The crucial element of this analysis lies in the fact that the above represents only half the story. Shifting our commuter to the other route puts him ahead of travelers who arrive after him on that route, making them worse off as a result of information.

The default departure rates on the two routes are in the ratio of the respective service rates (see Equations 4.5 and 4.6). For \( s_1 < s_2 \), the travel pattern on route 2 during the period where switching occurs from 2 to 1 consists of \( s_2/s_1 \) commuters\(^2\) whose default route is 2 alternating with a single commuter diverted from route 1, to form the travel rate of \( s_1 + s_2 \). Looking at the first such packet, the benefits to the switched commuter that we have just computed must be reduced by the increases in costs incurred by the \( s_2/s_1 \) commuters whose default route is 2 in order to obtain the net benefits to that packet. Let \( M = s_2/s_1 \). The first of the \( M \) commuters of the first packet who loses from switching incurs an extra travel time of \( 1/s_2 \) due to the traveler who switched and is now in front of her. The same applies to the remaining \( M - 1 \) travelers, and hence the total loss is given by \( \omega M/s_2 \). Therefore the net benefits to the first packet, \( B(t_a) \) are given by

\[ B(t_a) = \omega N'_A \ast \Delta s - \omega M \frac{1}{s_2} \]
\[ = \omega N'_A \ast \left( \frac{1}{s_a} - \frac{1}{s_1} \right) - \omega \frac{1}{s_1} \]  \hspace{1cm} (D.6)

\(^2\) Assume that \( s_2/s_1 \) is an integer for simplicity.
The first packet thus has positive net benefits if
\[
N'_A \cdot \left( \frac{1}{s_a} - \frac{1}{s_1} \right) > \frac{1}{s_1} \tag{D.7}
\]
Thus, there are positive net benefits to the first packet if the product of the number of incident-trapped people and the increase in service time on route 1 due to the incident is greater than the default service time on route 1. This condition is easily satisfied: If, for instance, the increase in service time is equal to the default service time (i.e. a 50% loss in capacity), then the condition is satisfied if the number of incident-trapped commuters is greater than one.

Next we determine the benefits for subsequent packets while switching to the incident-free route is still in progress. We begin by calculating the length of the switching phase.

The Length of the Switching Phase

For switching to occur, the potential costs on route 1 must be higher than on those on route 2. The potential costs of taking route 1 at time \( t_a + \Delta t \) can be computed in terms of offsets from the equilibrium cost \( C_{EQ} \). On route 1, no vehicles have passed \( B \) since \( t_a \). This contributes to a lowering of the cost relative to the equilibrium cost by an amount equal to the product of the number of commuters who have switched and their default service time on route 1. \( N'_A \) commuters will pass through the bottleneck at its reduced capacity, which adds to the relative cost by the product of the increase in the service time due to the incident and \( N'_A \). Thus the expected cost with info on route 1 at time \( t_a + \Delta t \) is given by\(^3\)

\[
C_1(t_a + \Delta t) = C_{EQ} + \omega[N'_A \Delta s - (s_1 \Delta t) \frac{1}{s_1}] \tag{D.8}
\]
\[
= C_{EQ} + \omega(N'_A \Delta s - \Delta t) \tag{D.9}
\]

The cost on route 2 at time \( t_a + \Delta t \) is given by

\[
C_2(t_a + \Delta t) = C_{EQ} + \omega(\frac{s_1}{s_2} \Delta t) \tag{D.10}
\]

Thus switching occurs as long as \( C_1(\Delta t) > C_2(\Delta t) \), i.e.

\[
N'_A \Delta s - \Delta t > \frac{s_1}{s_2} \Delta t \tag{D.11}
\]
or

\(^3\)Here we are making the implicit assumption that the queue length never falls to zero. Since switching will occur only until costs are equalized on the two routes, at a level higher than the original equilibrium cost, a zero queue would represent costs lower than \( C_{EQ} \), which is not possible.
\[ \Delta t < N_A' \frac{s_2}{s_1 + s_2} \Delta s \]

If the incident is long enough so that all of the \( N_T \) commuters pass through the bottleneck at reduced capacity, then the point in time at which switching ends is independent of the duration beyond that point in time at which the last of the trapped commuters passes through the bottleneck.

Suppose switching has just ended (for the first time), i.e.

\[ \Delta t = N_A' \frac{s_2}{s_1 + s_2} \Delta s \]

Will costs be stable such that switching will stop, as long as the original departure rates continue to occur? There are two cases to consider: The incident is “over”, in which case equilibrium rates will flow through equilibrium capacity bottlenecks. This means that there are zero diversions relative to the equilibrium costs, maintaining the current (higher) costs. In this case, attainment of this “pseudoequilibrium” also signifies permanent cessation of switching.

If the incident is still in progress, costs will rise on the incident route relative to the incident-free route. The information system will therefore need to occasionally switch vehicles to the incident-free route. If the costs are not equal when the incident ends, then switching will continue until costs are balanced, at which point it will stop, using the argument from the previous paragraph. The same argument says that if the costs happen to be equal as the incident ends, then switching will end immediately.

**Benefits for Subsequent Packets while Switching is in Progress**

Let us now look at the net benefits of a packet at time \( \Delta t \geq 0 \) after the beginning of the incident. Assume that we are still switching commuters from 1 to 2. two cases must be considered.

**Incident Still in Progress “at B”** The first case to be considered is that all of the commuters “in front” of the traveler whose default route is 1 and who is at \( B \) at time \( t_a + \Delta t \), will pass through the bottleneck at its reduced capacity. In this case we say the incident is “in progress at B” for clarity, but it means of course that the incident is in progress beyond \( t_a + \Delta t \). For a particular network / incident combination this case may not occur at all (This case implies that \( \Delta t \) as \( t_a + [N_T + s_1 \Delta t]/s_a \leq t_e \), or, as \( \Delta t \leq (N_A - N_T)/s_1 \). In this case \( N_A' = N_T \), or \( N_A \geq N_T \). In addition the switching assumption requires that \( \Delta t \leq N_T \frac{s_2}{s_1 + s_2} \Delta s \)

Again, to find benefits, we first compute the without and with information costs of the commuter who is switched, \( C_i^N(t_a + \Delta t) \).
Appendix D  How Information Results in Benefits

\[ C_1^N(t_a + \Delta t) = C_{EQ} + \omega (N_T + s_1 \Delta t) * \Delta s \]  \hspace{1cm} (D.14)

Also,

\[ C_1^Y(t_a) = C_{EQ} + \omega (s_1 \Delta t) \frac{1}{s_2} \]  \hspace{1cm} (D.15)

and thus the net benefits \( B_1(t_a + \Delta t) \) to the switcher at time \( t_a + \Delta t \) are

\[ B_1(t_a + \Delta t) = \omega [C_{EQ} + (N_T + s_1 \Delta t) * \Delta s] - \omega [C_{EQ} + (s_1 \Delta t) \frac{1}{s_2}] \]

\[ = \omega [(N_T + s_1 \Delta t) * \Delta s - (s_1 \Delta t) \frac{1}{s_2}] \]  \hspace{1cm} (D.16)

The losses for the \( M \) commuters on route 2 are caused by the delay due to the \( s_1 \Delta t \) commuters who have switched, and hence given by

\[ B_2(t_a + \Delta t) = -\omega \frac{s_2}{s_1} (s_1 \Delta t) \frac{1}{s_2} \]

\[ = -\omega \Delta t \]  \hspace{1cm} (D.17)

Hence net benefits \( B(t_a + \Delta t) \) for the packet passing \( B \) at time \( t_a + \Delta t \) are given by

\[ B(t_a + \Delta t) = \omega [(N_T + s_1 \Delta t) * \Delta s - (s_1 \Delta t) \frac{1}{s_2}] - \omega \Delta t \]

\[ = \omega [(N_T + s_1 \Delta t) * \Delta s - \frac{s_1}{s_2} \Delta t - \Delta t] \]

\[ = \omega [N_T \Delta s + \Delta t(s_1(1/s_a - 1/s_1) - \frac{s_1}{s_2} - 1)] \]

\[ = \omega [N_T \Delta s + \Delta t(s_1(1/s_a - 1/s_2) - 2)] \]  \hspace{1cm} (D.18)

Thus the net benefits accruing to a packet at \( \Delta t \) after the beginning of the incident are directly proportional to the product of the number of commuters trapped between \( B \) and \( C \) at the moment the incident occurred, and the increase in service time. Since \( s_2 > s_1 > s_a \) by assumption, the net benefits are a decreasing function of the time \( \Delta t \) only if \( s_1(1/s_a - 1/s_2) < 2 \). Hence benefits may increase or decrease with time after the beginning of the incident, depending on the relative magnitudes of the default capacities and the severity of the incident. In particular, for a given \( s_1 \), the more severe the incident, and the larger the capacity of the incident-free route, the greater the benefits.

Incident "Over" at \( B \) Now assume that not all the vehicles ahead of the commuter who is at \( B \) at time \( t_a + \Delta t \) will contribute to his added delay (in the no-information case), since
some will pass the bottleneck when it has returned to its normal operating capacity $s_1$.

The first case is when this situation holds for the first packet as well (i.e. when $\Delta t = 0$). In this case $N_A \leq N_T$ and $N'_A = N_A$. From above, $\Delta t$ is bounded simply by the switching limit, and hence $0 \leq \Delta t \leq N_A \frac{s_1 s_2}{s_1 + s_2} \Delta s$.

In the second case $N_A > N_T$ and $N'_A = N_T$, and thus the the incident "ends" at a time $\Delta t = [N_A - N_T]/s_1 > 0$. Therefore the limits for $\Delta t$ are given by $[N_A - N_T]/s_1 \leq \Delta t \leq N_T \frac{s_2}{s_1 + s_2} \Delta s$.

The added delay is simply a constant which equals the product of the duration of the incident and the ratio of the reduction in service rate to the original service rate. Therefore benefits are given by

$$B(t_a + \Delta t) = \omega [d \frac{s_1}{s_1} - \frac{s_a}{s_2} (s_1 \Delta t) \frac{1}{s_2} - \Delta t]$$

$$= \omega [d \frac{s_1}{s_1} - \frac{s_1 + s_2}{s_2} \Delta t]$$

(D.19)

Hence benefits for packets shifted during the period of redirection will decrease with time once the vehicles ahead of the commuter to be switched do not all pass through the bottleneck when it is experiencing reduced capacity - irrespective of the relative values of $s_1, s_2$, and $s_a$.

Benefits after Switching Stops

The final point to be investigated is the evolution of benefits after switching stops. Assume that the incident is over when switching ends, such that it will not resume (i.e. $\text{Deltat} > [s_a d - N_T]/s_1$). We substitute the value of $\Delta t$ from Equation D.13 in Equation D.19. There are two sub-cases, which have been discussed above. In the first $N'_A = N_A$, and hence we get

$$B(t_a + \Delta t) = \omega [d \frac{s_1}{s_1} - \frac{s_a}{s_2} N'_A \frac{s_2}{s_1 + s_2} \Delta s]$$

$$= \omega [d \frac{s_1}{s_1} - s_a d \Delta s]$$

$$= \omega [d \frac{s_1}{s_1} - ds_a (\frac{1}{s_a} - \frac{1}{s_1})]$$

$$= \omega [d \frac{s_1}{s_1} - d \frac{s_1 - s_a}{s_1}] = 0$$

(D.20)

Thus, there are zero net benefits to the first packet immediately after switching ends. This state of affairs will be maintained as long as the departures continue on both routes, since, as argued above, the costs will not change.
In the second case, $N_A' = N_T < N_A$, and hence the negative in the expression for the benefits is less than for the first case, implying that constant positive net benefits will accrue to the packet after switching stops.

### D.3 Analyzing Benefits: $s_1 > s_2$.

In this case, $s_1/s_2$ travelers are switched to route 2 for every commuter with default route 2.

**Benefits for First Packet**

The first traveler to be switched from 1 to 2 in the first packet will have the same with and without information costs as the first (and only) traveler to be switched when $s_1 < s_2$. The benefits are thus given by

$$\omega(C_1'(t_a) - C_1^N(t_a)) = \omega N_A' \Delta s$$  \hspace{1cm} (D.21)

The no-information costs for the next traveler to be switched will increase by the delay incurred by one traveler, namely $\Delta s$, and the information costs will increase by the extra delay imposed by the first switched commuter, $1/s_2$. This pattern repeats itself to yield benefits to the first packet that are given by

$$A = \sum_{i=1}^{s_1/s_2} [N_A' + (i - 1)]\Delta s - (i - 1)1/s_2$$  \hspace{1cm} (D.22)

The loss, $B$, to the single commuter with default route 2 is

$$B = \frac{s_1}{s_2} \frac{1}{s_2}$$  \hspace{1cm} (D.23)

Hence for positive benefits to the first packet the following must hold

$$\sum_{i=1}^{s_1/s_2} [N_A' + (i - 1)]\Delta s - (i - 1)1/s_2 - \frac{s_1}{s_2^2} > 0$$  \hspace{1cm} (D.24)

$$\sum_{i=0}^{s_1/s_2-1} [N_A' + i]\Delta s - i/s_2 - \frac{s_1}{s_2^2} > 0$$  \hspace{1cm} (D.25)

The formula for an arithmetic sum is given by

$$\sum_{k=0}^{n-1} (a + kr) = \frac{n}{2} [2a + (n - 1)r] = \frac{n}{2} (a + l)$$

where $l$ is the last term in the sum. The condition for positive benefits to the first packet can then be written as

$$\frac{s_1}{s_2} [2N_A'\Delta s + (\frac{s_1}{s_2} - 1)\Delta s - \frac{s_1 - s_2}{s_2^2}] > 0$$  \hspace{1cm} (D.26)
D.4 Conclusions

Benefits for Subsequent Packets while Switching is in Progress

The length of the switching phase depends only on the fact that vehicles are switched from 1 to 2 as long as costs are higher on 1, and therefore the expression for this quantity remains unaffected by the change in relative values of capacity. Benefits for subsequent packets must again be computed for two possible cases. We will modify the equations for the case $s_1 < s_2$, proceeding in a similar way to the just derived first packet case.

Incident Still in Progress “at B” The net benefits $B_1(t_a + \Delta t)$ to the $M$ switchers at time $t_a + \Delta t$ are

$$B_1(t_a + \Delta t) = \omega \sum_{i=1}^{s_1/s_2} [N_T + s_1 \Delta t + (i - 1)] \Delta s - [s_1 \Delta t + (i - 1)] 1/s_2$$

(D.27)

The loss for the commuter on route 2 is caused by the delay due to the $s_1 \Delta t + s_1/s_2$ commuters who have switched, and equals

$$B_2(t_a + \Delta t) = -\omega(s_1 \Delta t + \frac{s_1}{s_2}) \frac{1}{s_2}$$

(D.28)

Net benefits in this case are given by

$$B(t_a + \Delta t) = \omega \sum_{i=1}^{s_1/s_2} [N_T + s_1 \Delta t + (i - 1)] \Delta s - [s_1 \Delta t + (i - 1)] 1/s_2 - \omega(s_1 \Delta t + \frac{s_1}{s_2}) \frac{1}{s_2}$$

(D.29)

Incident “over” at B Let $\Phi = d \frac{t_a - t_a}{s_1}$. Net benefits in this case are given by

$$B(t_a + \Delta t) = \omega \sum_{i=1}^{s_1/s_2} [\Phi - (s_1 + [i - 1])] - (s_1 \Delta t + \frac{s_1}{s_2}) \frac{1}{s_2}$$

(D.30)

Benefits after Switching Stops

Assume that the incident is over when switching ends, such that it will not resume.

$$B(t_a + \Delta t) = \omega \sum_{i=1}^{s_1/s_2} [\Phi - (s_1 + [i - 1])] - (s_1(N_4' + \frac{s_2}{s_1 + s_2} \Delta s) + \frac{s_1}{s_2}) \frac{1}{s_2}$$

(D.31)

D.4 Conclusions

We have looked at how benefits are generated when a perfect information system guides vehicles to the shortest route in response to an incident. To capture the fact that there are both losers and gainers as a result of the route switching, we determined benefits for
packets that are composed of the travelers on both routes that pass the information node $B$ during a fixed time period.

The incident was assumed on route 1. We examined the case in which $s_1 < s_2$, and the case in which $s_1 > s_2$. The latter case lead to more involved expressions that do not generalize well without making some assumptions about the relative magnitudes of the default and incident capacities. Conditions were derived for positive benefits for packets at three different times: (1) Just after the beginning of the incident, (2) At an arbitrary time after the incident when switching is still in progress, and (3) after switching stops. There are several possible sub-cases that may occur for the latter two times.

The following observations can be made when $s_1 < s_2$. Benefits to the initial packet are positive under easily satisfiable conditions. During the switching phase benefits may increase or decrease depending on the relative magnitudes of the capacities. After switching, benefits may remain at zero, or be positive. When $s_1 > s_2$, more complex expressions arise that are not easily interpreted without making more concrete assumptions about the relative magnitudes of the capacities.

Even for a simple case as the one considered here, numerous cases must be considered, and the results are not always very simple. In future work, we intend to further analyze this situation in order to draw more general conclusions.